

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2005, 10:05:19 ; Search time 169 Seconds  
(without alignments)  
2213.002 Million cell updates/sec

Title: US-09-989-687-126  
Perfect score: 5287  
Sequence: 1 MRAVPEGFRKRLGSDMGN.....CDPLKKPKHFIDFTMAECS 967

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

\* Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5232	99.0	967	2 AAW80285	AAW80285 Human int
2	5224	98.8	967	2 AAY04142	Aay04142 Human Tan
3	5224	98.8	967	8 ADR14133	Adr14133 Human NP-
4	5224	98.8	967	8 ADQ39940	Adq39940 Human myo
5	5224	98.8	967	8 ADQ39941	Adq39941 Human myo
6	5221.5	98.8	968	4 AAB50011	Aab50011 Protein;
7	5201	98.4	967	2 AAW78189	Aaw78189 Human sec
8	5201	98.4	967	6 ADA57139	Ada57139 Human sec
9	5201	98.4	967	6 ADA41003	Ada41003 Human sec
10	5201	98.4	967	7 ADB91631	Adb91631 Human sec
11	5201	98.4	967	7 ADC74267	Adc74267 Human sec
12	5201	98.4	967	7 ADD37948	Add37948 Human sec
13	5144	97.3	950	2 AAY49501	Aay49501 Human MET
14	5144	97.3	950	4 AAB73549	Aab73549 Human ADA
15	5144	97.3	950	4 AAB50002	Aab50002 Human MET
16	5140	97.2	949	7 ADG72483	Adg72483 Human agg
17	5136	97.1	950	8 ADQ20218	Adq20218 Human PRO
18	5136	97.1	950	8 ADQ39942	Adq39942 Human myo
19	4293.5	79.7	968	6 ABU08387	Abu08387 Murine ma
20	4215.5	79.7	950	3 AAY53899	Aay53899 Amino aci
21	3996.5	75.6	896	3 AAB21265	Aab21265 Mouse met
22	3922	74.2	727	2 AAW78435	Aaw78435 Human ADA
23	2495.5	47.2	924	5 ABP70062	Abp70062 Human NOV
24	2477.5	46.9	950	5 AAE22541	Aae22541 Human pro
25	2477.5	46.9	950	6 ABU09520	Abu09520 Human pro

26	2477.5	46.9	950	8 ADQ88214	Adq88214 Human 655
27	2473.5	46.8	950	6 ABR40092	Abx40092 Human ADA
28	2471.5	46.7	950	4 AAG62299	Aag62299 Human met
29	2455.5	46.4	952	5 AAU74751	Aau74751 Human pro
30	2408	45.5	928	5 AAU72899	Aau72899 Human met
31	2323.5	43.9	890	6 ABP96306	Abp96306 Human ADA
32	2321.5	43.9	890	8 ADK70513	Adk70513 Respirato
33	2320.5	43.9	890	2 AAY49502	Aay49502 Human MET
34	2320.5	43.9	890	4 AAB50003	Aab50003 Human MET
35	2313	43.7	889	4 AAB74946	Aab74946 Human ADA
36	2265.5	42.9	905	4 AAB72284	Aab72284 Murine AD
37	2190	41.4	823	6 ABU08383	Abu08383 Human met
38	2129	40.3	837	4 AAG78228	Aag78228 Human egg
39	2129	40.3	837	7 ADB85488	Adb85488 Human egg
40	2129	40.3	840	3 AAB21256	Aab21256 Human met
41	2124	40.2	837	2 AAW75425	Aaw75425 Human agg
42	2124	40.2	837	7 ADJ69542	Adj69542 Human hea
43	2124	40.2	837	8 ADS20209	Ads20209 Human agg
44	2121	40.1	846	8 ADS20232	Ads20232 Human agg
45	2117	40.0	837	3 AAY99429	Aay99429 Human PRO

## ALIGNMENTS

RESULT 1  
AAW80285  
ID AAW80285 standard; protein; 967 AA.  
XX  
AC AAW80285;  
XX  
DT 19-JAN-1999 (first entry)  
XX  
DE Human integrin ligand polypeptide ITGL-TSP.  
XX

KW ITGL-TSP; integrin ligand; angiogenic disease; cancer; atherosclerosis;  
KW Chronic inflammatory disorder; rheumatoid arthritis; tissue remodeling;  
KW macular degeneration; diabetic retinopathy; Alzheimer's disease; human;  
KW restenosis.

XX Homo sapiens.

XX EP874050-A2.

XX 28-OCT-1998.

XX 27-JAN-1998; 98EP-00300575.

XX 24-APR-1997; 97US-00845496.

XX (SMIK ) SMITHKLINE BEECHAM CORP.

XX (SMIK ) SMITHKLINE BEECHAM PLC.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Jonak ZL, Trulli SH, Fronwald JA, Hastings GA, Terrett JA;

XX WPI; 1998-544643/47.

XX N-PSDB; AAV66508.

XX DNA encoding integrin ligand polypeptide ITGL-TSP - used to treat  
XX angiogenic diseases, restenosis, Alzheimer's disease and in tissue  
XX remodeling.

XX Claim 11; Page 6-9; 24pp; English.

XX This represents a human integrin ligand polypeptide ITGL-TSP. ITGL-TSP  
XX polypeptides can be used in the treatment of angiogenic diseases such as  
XX cancer, cancer metastasis, chronic inflammatory disorders, rheumatoid  
XX arthritis, atherosclerosis, macular degeneration or diabetic retinopathy,  
XX restenosis, Alzheimer's disease and tissue remodeling. They can be used  
XX to treat a subject in need of enhanced activity or expression of the ITGL  
XX -TSP polypeptide

SQ	Sequence 967 AA;	
Query Match	99.0%; Score 5232; DB 2; Length 967;	
Best Local Similarity	99.0%; Pred. No. 0;	
Matches 957; Conservative	4; Mismatches 6; Indels 0; Gaps 0;	
Qy	1 MRAVPEGFRKLGSDMGNAERAPGSRSGFVPPTLLLLAAALLAVSDALGRPSEDEEL 60	
Db	1 MRAVPEGFRKLGSDMGNAERAPGSRSGFVPPTLLLLAAALLAVSDALGRPSEDEEL 60	
Qy	61 VPELERVPGHGTTRLRHLHAFDQQLDLVDPDSFLAPGFTLQNVGRKSGSDTLPETDL 120	
Db	61 VPELERAPGHGTTRLRHLHAFDQQLDLVDPDSFLAPGFTLQNVGRKSGSETPLPETDL 120	
Qy	121 AHCFYSGTVNGDPSSAAALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAP 180	
Db	121 AHCFYSGTVNGDPSSAAALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAP 180	
Qy	181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETDEDEGTEGEGPQWSPQDPALQGVG 240	
Db	181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETDEDEGTEGEGPQWSPQDPALQGVG 240	
Qy	241 OPTGTGSIKRRKRFVSSHRYVETMLVADQSMAEFHGSLKHLLTLFSAARLYKHPISRN 300	
Db	241 OPTGTGSIKRRKRFVSSHRYVETMLVADQSMAEFHGSLKHLLTLFSAARLYKHPISRN 300	
Qy	301 SVSLVWVKILVIHDEQKPEVTSNAALTLRNCWQKQHNPPSRDABHYDTAILFTROD 360	
Db	301 SVSLVWVKILVIHDEQKPEVTSNAALTLRNCWQKQHNPPSRDABHYDTAILFTROD 360	
Qy	361 LCGSQTCDTLGADVTCVDPSSRSVLEDDGLQAAFTTAHELGHVFMNPHDDAKQACSL 420	
Db	361 LCGSQTCDTLGADVTCVDPSSRSVLEDDGLQAAFTTAHELGHVFMNPHDDAKQACSL 420	
Qy	421 NGVNQDSHMAMSLNLDHSPWSPCSGYMTITSFLDNGHGCLMDKPNPIQLPDLPGT 480	
Db	421 NGVNQDSHMAMSLNLDHSPWSPCSGYMTITSFLDNGHGCLMDKPNPIQLPDLPGT 480	
Qy	481 SYDANROCOFTFGEDSKHCPDAASTCTLWCTGTSGGVILVCOTKHPFWDGTSCEGKWC 540	
Db	481 SYDANROCOFTFGEDSKHCPDAASTCTLWCTGTSGGVILVCOTKHPFWDGTSCEGKWC 540	
Qy	541 INGCNVNKHNRKHFDTPPHGSGWGMWPGDCSRTCCTGGGVQVYTMRECDNPVPKNGKCEG 600	
Db	541 INGCNVNKHNRKHFDTPPHGSGWGMWPGDCSRTCCTGGGVQVYTMRECDNPVPKNGKCEG 600	
Qy	601 KRVRYSNLEDCPDNNGKTFREOCEAHNEFSAFSGSGPAVEWIPKYAGVSPKDRCKL 660	
Db	601 KRVRYSNLEDCPDNNGKTFREOCEAHNEFSAFSGSGPAVEWIPKYAGVSPKDRCKL 660	
Qy	661 ICQAKGTGYFVLQPKVVDGTPCSPDSTSVQCGQVKGAGCDRIIDSKKFKDKGVCNGN 720	
Db	661 ICQAKGTGYFVLQPKVVDGTPCSPDSTSVQCGQVKGAGCDRIIDSKKFKDKGVCNGN 720	
Qy	721 GSTCKKTSVTSAPGVHDIITITPTGATNIEVKORNGSRNNGSFLAIKAAADGTVILN 780	
Db	721 GSTCKKTSVTSAPGVHDIITITPTGATNIEVKORNGSRNNGSFLAIKAAADGTVILN 780	
Qy	781 GDYTLSTLEQDQIMYKGVVLYRSGSSAALERIRSFSPKLEPTIIOVLTVGNALRPKIYTY 840	
Db	781 GDYTLSTLEQDQIMYKGVVLYRSGSSAALERIRSFSPKLEPTIIOVLTVGNALRPKIYTY 840	
Qy	841 FVKKKESFNAIPTFSAWVIEWECSEKSCBLGHORRLVECRDINGOPASECAKEVPAS 900	
Db	841 FVKKKESFNAIPTFSAWVIEWECSEKSCBLGHORRLVECRDINGOPASECAKEVPAS 900	
Qy	901 TRPCADHPCPQWQLGEWSSCKTCGKYKTKSLKCLSHDGGVLSHSDSCDPLKKPKHFDIF 960	
Db	901 TRPCADHPCPQWQLGEWSSCKTCGKYKTKSLKCLSHDGGVLSHSDSCDPLKKPKHFDIF 960	
Qy	961 CTMAECS 967	
Db	961 CTMAECS 967	

RESULT 2		
AA04142		
ID	AA04142 standard; protein; 967 AA.	
XX	AC	AA04142;
XX	DT	15-JUN-1999 (first entry)
XX	DE	Human Tango-71 protein.
XX	KW	Human; Tango-71; Tango-73; Tango-74; Tango-76; Tango-83; diagnosis; detection.
XX	OS	Homo sapiens.
XX	PN	WO9907850-A1.
XX	PD	18-FEB-1999.
XX	PF	06-AUG-1998; 98WO-US016502.
XX	PR	06-AUG-1997; 97US-0054966P.
XX	PR	05-SEP-1997; 97US-0058108P.
XX	PA	(MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX	PI	Holtzman DA, Goodearl ADJ;
XX	DR	WPI; 1999-167426/14.
XX	DR	N-PSDB; AAX19955.
XX	PT	New TANGO polypeptides and nucleic acids encoding them - useful as diagnostic agents and for treating disorders caused by aberrant expression of TANGO.
XX	PS	Claim 8; Fig 1; 84pp; English.
XX	CC	The present sequence represents human Tango-71. Tango polypeptides are useful for identifying compounds which bind the polypeptide via direct binding, competition binding assays or Tango-71, -73, -74, 76 or -83 mediated signal transduction. Tango polypeptides are also useful for identifying modulating compounds by determining effect on Tango activity.
XX	CC	Tango polypeptides and nucleic acids are useful for diagnosing diseases related to aberrant expression of Tango, and Tango polypeptides are useful for raising antibodies which can be used in diagnostic assays for detection of Tango, and also for generating anti-idiotypic antibodies for prevention and protection
XX	SQ	Sequence 967 AA;
Query Match	98.8%; Score 5224; DB 2; Length 967;	
Best Local Similarity	98.9%; Pred. No. 0;	
Matches 956; Conservative	4; Mismatches 7; Indels 0; Gaps 0;	
Qy	1 MRAVPEGFRKLGSDMGNAERAPGSRSGFVPPTLLLLAAALLAVSDALGRPSEDEEL 60	
Db	1 MRAVPEGFRKLGSDMGNAERAPGSRSGFVPPTLLLLAAALLAVSDALGRPSEDEEL 60	
Qy	61 VPELERVPGHGTTRLRHLHAFDQQLDLVDPDSFLAPGFTLQNVGRKSGSDTLPETDL 120	
Db	61 VPELERAPGHGTTRLRHLHAFDQQLDLVDPDSFLAPGFTLQNVGRKSGSETPLPETDL 120	
Qy	121 AHCFYSGTVNGDPSSAAALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAP 180	
Db	121 AHCFYSGTVNGDPSSAAALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAP 180	
Qy	181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETDEDEGTEGEGPQWSPQDPALQGVG 240	
Db	181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETDEDEGTEGEGPQWSPQDPALQGVG 240	
Qy	241 OPTGTGSIKRRKRFVSSHRYVETMLVADQSMAEFHGSLKHLLTLFSAARLYKHPISRN 300	

```
Db 241 QPTGTSIRKRFVSHRYVETMLVADQSMAEFHGSLKHYLLTLTFSVAARLYKPSIRN 300
Qy 301 SVSLVVKLVTHDQKGPVTSNAALTLRNFQKQKQHPSPDRDAEHYDTAILFTTQD 360
Db 301 SVSLVVKLVTHDQKGPVTSNAALTLRNFQKQKQHPSPDRDAEHYDTAILFTTQD 360
Qy 361 LCGSQTCDTLGMADVGTVCDPSSRSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
Db 361 LCGSQTCDTLGMADVGTVCDPSSRSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
Qy 421 NGVNDSHMMASMLNLSLHDSQWPSFCGMYMTSFLDNGHGECLMDKPNPIQLPGDLPGT 480
Db 421 NGVNDSHMMASMLNLSLHDSQWPSFCGMYMTSFLDNGHGECLMDKPNPIQLPGDLPGT 480
Qy 481 SYDANRQCOFTGEGDSKCPDAASTCTLWCTGTGSGVLVCOVKHPPWADGTSCGEGKWC 540
Db 481 SYDANRQCOFTGEGDSKCPDAASTCTLWCTGTGSGVLVCOVKHPPWADGTSCGEGKWC 540
Qy 541 INGKCVNKNHRKHFDTPFHGSGMGMGPMGDCSRTCGGVQVYTMRECDNPVPKNGGKYCEG 600
Db 541 INGKCVNKTDRKHFDTPFHGSGMGMGPMGDCSRTCGGVQVYTMRECDNPVPKNGGKYCEG 600
Qy 601 KVRVRSNLEDCPNNGKTFREOCEAHNEFSKASFGSGPAVEWIPKYAGVSPKORCKL 660
Db 601 KVRVRSNLEDCPNNGKTFREOCEAHNEFSKASFGSGPAVEWIPKYAGVSPKORCKL 660
Qy 661 ICQAKGIGVFFVLQPKVVDGTPCSPDSTSVQGVQCVKAGCDRIIDSKKPKDCKGCGGN 720
Db 661 ICQAKGIGVFFVLQPKVVDGTPCSPDSTSVQGVQCVKAGCDRIIDSKKPKDCKGCGGN 720
Qy 721 GSTCKKISGSVTSAPGVDHIIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTIYL 780
Db 721 GSTCKKISGSVTSAPGVDHIIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTIYL 780
Qy 781 GDYTLSTLEQDIWYKGVVLRYSGSSAALIRISFPLKEPLTIQVLTGNALRPKIKTY 840
Db 781 GDYTLSTLEQDIWYKGVVLRYSGSSAALIRISFPLKEPLTIQVLTGNALRPKIKTY 840
Qy 841 FVKKKESFNALPTFSAMVIBEWGSCSKCELGWQRRVCECRDINGQPASECAKEVKPAS 900
Db 841 FVKKKESFNALPTFSAMVIBEWGSCSKCELGWQRRVCECRDINGQPASECAKEVKPAS 900
Qy 901 TRPCADHPCPQWLGEWSSCSKTCGKYKTKSLKCLSHDGGVLSHSDCPLKKPKHFIDF 960
Db 901 TRPCADHPCPQWLGEWSSCSKTCGKYKTKSLKCLSHDGGVLSHSDCPLKKPKHFIDF 960
Qy 961 CTMAECS 967
Db 961 CTMAECS 967

RESULT 3
ADRI4133
ID ADRI4133 standard; protein; 967 AA.
AC ADRI4133;
XX
XX
DT 21-OCT-2004 (first entry)
DE
DE Human NF-kappaB pathway-associated protein SeqID134.
KW NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
KW antiarthritic; antiarthritis; gastrointestinal-Gen; antiasthmatic;
KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
KW immunosuppressive; vulnary; gene therapy; immune disorder;
KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant
KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
KW viral replication; host cell survival; evasion of immune response;
KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
```

```
KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
KW autoimmune disorder; hyper immune activity;
KW aberrant acute phase response; hypercongenital condition; birth defect;
KW necrotic lesion; wound; organ transplant rejection;
KW aberrant signal transduction; proliferating disorder; cancer;
KW HIV propagation; human.
XX
OS Homo sapiens.
XX
PN WO2004065577-A2.
XX
PD 05-AUG-2004.
XX
PF 13-JAN-2004; 2004WO-US000798.
PR 14-JAN-2003; 2003US-0440068P.
PR 12-MAY-2003; 2003US-0469757P.
XX
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Nadler SG, Neubauer MG, Feder JN, Carman J;
XX
XX WPI: 2004-562168/54.
XX
XX N-PSDB; ADRI4132.
XX
XX New isolated polynucleotides and polypeptides associated with NF-kappaB
XX pathway, useful for diagnosing, treating, or preventing disorders or
XX diseases associated with NF-kappaB pathway.
XX
XX Claim 6; SEQ ID NO 134; 237pp; English.
XX
CC This invention relates to the novel association of protein sequences (and
CC the genes which encode them) to the NF-kappaB pathway. The invention may
CC be useful for the production of compounds with an antiinflammatory,
CC cytostatic, hepatotropic, virucide, antiarthritic, antiarthritis,
CC gastrointestinal-Gen, antiasthmatic, antiarteriosclerotic,
CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
CC vulnary activity or for gene therapy. The proteins and nucleotides are
CC useful for diagnosing, preventing, treating, or ameliorating conditions
CC or diseases associated with the NF-kappaB pathway. The condition is an
CC immune disorder, an inflammatory disorder, an inflammatory disorder
CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM
CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory
CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
CC immune activity, disorders related to aberrant acute phase responses,
CC hypercongenital conditions, birth defects, necrotic lesions, wounds,
CC organ transplant rejection, conditions related to organ transplant
CC rejection, disorders related to aberrant signal transduction.
CC proliferating disorders, cancers and HIV propagation in cells infected
CC with other viruses. The present sequence is that of a human protein which
CC is subject to the novel association with the NF-kappaB pathway of the
CC invention. Note: This sequence does not appear in the specification but
CC was obtained by the indexer from Genbank.
XX
SQ Sequence 967 AA;
```

Query Match 98.8%; Score 5224; DB 8; Length 967;

Best Local Similarity 98.8%; Pred. NO. 0;

Matches 955; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MQRVPEGFGRKLGSDMGNAERAPSGSFVPVTLTLLAALAVSDALGRPSEDEL 60

Db 1 MQRVPEGFGRKLGSDMGNAERAPSGSFVPVTLTLLAALAVSDALGRPSEDEL 60

Qy 61 VVPELERYPGHGTTRLRHLHAFDQQLDLDVDPSSFLAPGFTLQNVGRKSGSDTLPETDL 120

Db 61 VVPELERYPGHGTTRLRHLHAFDQQLDLDVDPSSFLAPGFTLQNVGRKSGSDTLPETDL 120

QY 121 AHCFYSGTVNGDPSSAAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKPPAP 180  
Db |||||  
QY 121 AHCFYSGTVNGDPSSAAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKPPAP 180  
Db |||||  
QY 181 LQFHLLRRNRQGVGTCGVVDDEPRPTGKAFTEDDEDEGEDEGPGQWSPQDPALQGVG 240  
Db |||||  
QY 241 OPTGTGSTRKRKFVSSHRYVETMLVADQSMABFHGSGLKHLLTLFVSAARLYKHPSIRN 300  
Db |||||  
QY 241 OPTGTGSTRKRKFVSSHRYVETMLVADQSMABFHGSGLKHLLTLFVSAARLYKHPSIRN 300  
Db |||||  
QY 301 SVSLVVKILVTHDQKGPVTSNAALTLRNFCKWQKOHNPSPDRDAEHYDTAILFTQD 360  
Db |||||  
QY 301 SVSLVVKILVTHDQKGPVTSNAALTLRNFCKWQKOHNPSPDRDAEHYDTAILFTQD 360  
Db |||||  
QY 361 LCGSQTCDTLGMADVTCDFPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDADKQACASL 420  
Db |||||  
QY 361 LCGSQTCDTLGMADVTCDFPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDADKQACASL 420  
Db |||||  
QY 421 NGVNQDSHMAASMLNLDHSPQWSPSCGYMITSFLDNGHGCLMDKPNPIQLPGLDPT 480  
Db |||||  
QY 421 NGVNQDSHMAASMLNLDHSPQWSPSCGYMITSFLDNGHGCLMDKPNPIQLPGLDPT 480  
Db |||||  
QY 481 SYDANRQCOFTFGEDSKHCPDAASTCTLWCTGTSGGVLVCQTKHFPWADGTSCEGKWC 540  
Db |||||  
QY 481 SYDANRQCOFTFGEDSKHCPDAASTCTLWCTGTSGGVLVCQTKHFPWADGTSCEGKWC 540  
Db |||||  
QY 541 INGCNVNKHRRKHFDTPPHGSGMWGMPGDCSRTCGGQVQVYTRMRECDNPVPKNGKCEG 600  
Db |||||  
QY 541 INGCNVNKHRRKHFDTPPHGSGMWGMPGDCSRTCGGQVQVYTRMRECDNPVPKNGKCEG 600  
Db |||||  
QY 601 KRVYRSCNLEDCPDNNGKTFREOCEAHNFEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660  
Db |||||  
QY 601 KRVYRSCNLEDCPDNNGKTFREOCEAHNFEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660  
Db |||||  
QY 661 ICOAKGIGYFFVLPQKVVDGTPCSPDSTSVCGQCVKAGCDRIIDSKKFPDKGVCQGN 720  
Db |||||  
QY 661 ICOAKGIGYFFVLPQKVVDGTPCSPDSTSVCGQCVKAGCDRIIDSKKFPDKGVCQGN 720  
Db |||||  
QY 721 GSTCKKISGVSATKPGYHDIITPTGATNIEVKORNGSRNNGSFLAIKAADGTYLN 780  
Db |||||  
QY 721 GSTCKKISGVSATKPGYHDIITPTGATNIEVKORNGSRNNGSFLAIKAADGTYLN 780  
Db |||||  
QY 781 GDYTLSTLEQDIMYKGVVLRYSGSSAALERIRSFSPKLEPTIOVLTGVGNALRPKIKTY 840  
Db |||||  
QY 781 GDYTLSTLEQDIMYKGVVLRYSGSSAALERIRSFSPKLEPTIOVLTGVGNALRPKIKTY 840  
Db |||||  
QY 841 FVKKKESFNAIPTFSAMVIEWGECSKSCBLGWQRRLVECRDINGQPASECAKEVPAS 900  
Db |||||  
QY 841 FVKKKESFNAIPTFSAMVIEWGECSKSCBLGWQRRLVECRDINGQPASECAKEVPAS 900  
Db |||||  
QY 901 TRPCADHPCPQWQGEWSSCKTCGKYKTSLSKLSHDGVLGSHDSCDPLKPKPHFIDF 960  
Db |||||  
QY 901 TRPCADHPCPQWQGEWSSCKTCGKYKTSLSKLSHDGVLGSHDSCDPLKPKPHFIDF 960  
Db |||||  
QY 961 CTMAECS 967  
Db |||||  
QY 961 CTMAECS 967  
Db |||||

RESULT 4

ID ADQ39940 standard; protein; 967 AA.  
XX  
AC ADQ39940;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human myocardial infarction-associated gene derived protein, SEQ ID 1603.  
XX  
KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;  
KW cardiant; gene therapy; human.

XX Homo sapiens.  
OS  
XX WO2004058052-A2.  
PN  
XX 15-JUL-2004.  
PD  
XX  
PF 22-DEC-2003; 2003WO-US040978.  
XX  
XX 20-DEC-2002; 2002US-0434778P.  
PR 10-MAR-2003; 2003US-0453135P.  
PR 30-APR-2003; 2003US-0466412P.  
PR 23-SEP-2003; 2003US-0504955P.  
XX  
XX (APPL-) APPLERA CORP.  
FA  
XX  
XX Cargill M, Devlin JJ, Iakoubova O;  
PI  
XX WPI: 2004-533949/51.  
XX  
DR N-PSDB, ADQ39112.  
XX  
XX  
PT Identifying an individual who has an altered risk for developing  
PT myocardial infarction by detecting a single nucleotide polymorphism in  
PT the individual's nucleic acids.  
XX  
XX Claim 10; SEQ ID NO 1603; 145pp; English.  
XX  
XX The invention relates to a novel method for identifying an individual who  
XX has an altered risk for developing myocardial infarction. The method  
XX comprises detecting a single nucleotide polymorphism (SNP) in any one of  
XX the nucleotide sequences given in the specification in the individual's  
XX nucleic acids, where the presence of the SNP is correlated with an  
XX altered risk for myocardial infarction in the individual. The invention  
XX further comprises: an isolated nucleic acid molecule comprising at least  
XX 8 contiguous nucleotides where one of the nucleotides is an SNP given in  
XX the specification or its complement and encoding any one of the amino  
XX acid sequences given in the specification; an isolated polypeptide  
XX comprising an amino acid sequence given in the specification; an antibody  
XX that specifically binds to the polypeptide or its antigen-binding  
XX fragment; an amplified polynucleotide containing an SNP given in the  
XX specification and which is between about 16 and 1000 nucleotides in  
XX length; a kit for detecting an SNP in a nucleic acid, comprising the  
XX polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a  
XX nucleic acid molecule; a method of detecting a variant polypeptide; and a  
XX method for identifying an agent useful in treating or preventing  
XX myocardial infarction. The novel detection method has cardiac activity.  
XX The nucleic acids of the invention may be used in gene therapy. The  
XX method is useful in identifying an individual who has an increased or  
XX decreased risk for developing myocardial infarction and for preparing a  
XX composition for treating or preventing myocardial infarction. This  
XX sequence represents the protein of a human myocardial infarction-  
XX associated gene containing one or more SNPs of the invention. Note: This  
XX sequence was not shown in the specification. The sequence has come from  
XX an electronic sequence listing downloaded from the WIPO website.

Sequence 967 AA;

Query Match 98.8%; Score 5224; DB 8; Length 967;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 956; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MQRVPEGFGRRKLGSDMGNAERAPGSRSGPVPPTLLLLAAALAVSDALGRSEDEEL 60  
Db |||||  
Qy 1 MQRVPEGFGRRKLGSDMGNAERAPGSRSGPVPPTLLLLAAALAVSDALGRSEDEEL 60  
Db |||||  
Qy 61 VVPELERVPGHGTTRLRHLHAFDQDLDDVPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120  
Db |||||  
Qy 61 VVPELERVPGHGTTRLRHLHAFDQDLDDVPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120  
Db |||||  
Qy 121 AHCFYSGTVNGDPSSAAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKPPAP 180  
Db |||||  
Qy 121 AHCFYSGTVNGDPSSAAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKPPAP 180  
Db |||||

Qy	181	LQFHLLRRNRQDVGCGT	CGVVDDEPRPKA	TEDEDETEGTEGDEGEPQ	WSPDPALQGVG	240			
Db	181	LQFHLLRRNRQDVGCGT	CGVVDDEPRPKA	TEDEDETEGTEGDEGEPQ	WSPDPALQGVG	240			
Qy	241	OPTGTGSRKRKRFVSSHR	VYETMLVADQSM	AEFHGSLKHYLLT	LFSVAARLYKHPSIRN	300			
Db	241	OPTGTGSRKRKRFVSSHR	VYETMLVADQSM	AEFHGSLKHYLLT	LFSVAARLYKHPSIRN	300			
Qy	301	SVSLVVVKILVHDHQBQ	PEVTSNAALTR	NFCNWKQHNPP	SDRDAEHYDTAILFT	TRQD 360			
Db	301	SVSLVVVKILVHDHQBQ	PEVTSNAALTR	NFCNWKQHNPP	SDRDAEHYDTAILFT	TRQD 360			
Qy	361	LCGSQTCDTLGNADVGT	CDPSRSCSV	EDDGLQAAFTTA	HELGHVFNPHDQAKQ	CASL 420			
Db	361	LCGSQTCDTLGNADVGT	CDPSRSCSV	EDDGLQAAFTTA	HELGHVFNPHDQAKQ	CASL 420			
Qy	421	NGVNQDSHMA	SMLSNLN	DHSPWSPCSG	YMITSLDNGHGECL	MDKPQNPIQLPGDL	PCT 480		
Db	421	NGVNQDSHMA	SMLSNLN	DHSPWSPCSAY	MITSLDNGHGECL	MDKPQNPIQLPGDL	PCT 480		
Qy	481	SYDANRQCQFTFG	EDSKHCPDAAS	TSTLWCTGTSG	GVLCVCTKHFP	WADGTSCGEGKWC	540		
Db	481	SYDANRQCQFTFG	EDSKHCPDAAS	TSTLWCTGTSG	GVLCVCTKHFP	WADGTSCGEGKWC	540		
Qy	541	INGKCVNKNH	RKHFDTP	PHGSGWGMW	PGWDCSRTC	CGGVQVYTMRECDN	PVPKXNGKCY	CEG 600	
Db	541	INGKCVNKTDR	KRKHFDTP	PHGSGWGMW	PGWDCSRTC	CGGVQVYTMRECDN	PVPKXNGKCY	CEG 600	
Qy	601	KRVYRSCNLEDC	PDNNNGKTF	REEQCEAHN	EFSKASFGSG	PAVEIPIKYAGV	SPKDRCKL 660		
Db	601	KRVYRSCNLEDC	PDNNNGKTF	REEQCEAHN	EFSKASFGSG	PAVEIPIKYAGV	SPKDRCKL 660		
Qy	661	ICQAKGIGYFF	VLQPKVVDG	TFPCSPDST	SVCVQGC	VKAGCDRIIDSK	KKPKCGV	CCGN 720	
Db	661	ICQAKGIGYFF	VLQPKVVDG	TFPCSPDST	SVCVQGC	VKAGCDRIIDSK	KKPKCGV	CCGN 720	
Qy	721	GSTCKKISG	SVTSAPKGY	YHDIITPTG	ATNTEVKORN	ORSGRNNSFLAI	KAADGT	TYILN 780	
Db	721	GSTCKKISG	SVTSAPKGY	YHDIITPTG	ATNTEVKORN	ORSGRNNSFLAI	KAADGT	TYILN 780	
Qy	781	GDYTLSTLEQ	DIMYKGVV	LYSGSAA	LERIRSF	SPKLEPTIQVLT	VGNALRP	KIKYTY 840	
Db	781	GDYTLSTLEQ	DIMYKGVV	LYSGSAA	LERIRSF	SPKLEPTIQVLT	VGNALRP	KIKYTY 840	
Qy	841	FVKKKKGS	FNAIP	TFSAWV	IEWGCSK	SELGWORRL	VECDINGQ	PASECAK	EVKPKAS 900
Db	841	FVKKKKGS	FNAIP	TFSAWV	IEWGCSK	SELGWORRL	VECDINGQ	PASECAK	EVKPKAS 900
Qy	901	TRPCADHP	CPQWLGEW	SSCKTCG	KGYKKTSL	CKLSHDGG	VLSHDS	CDPLKPK	KHFIDF 960
Db	901	TRPCADHP	CPQWLGEW	SSCKTCG	KGYKKTSL	CKLSHDGG	VLSHDS	CDPLKPK	KHFIDF 960
Qy	961	CTMAECS	967						
Db	961	CTMAECS	967						

RESULT 5  
AD039941

[illegible]



Db 181 PLQFHLRRNRQDVGTCGVVDDEPRPTGKAEETDEDEGTGEDEGPQWSPQDPALQGV 240  
QY 240 GOPTGTGSIKKRKFVSSHRYVETMLVADQSMAEFHGSLGKHYLLTLFSAARLYKHPsir 299  
Db 241 GOPTGTGSIKKRKFVSSHRYVETMLVADQSMAEFHGSLGKHYLLTLFSAARLYKHPsir 300  
QY 300 NSVSLVWVKILVIHDEQKPEVTSNAALTNRNFCWQKHNPSPDRDAEHYDTAILFTRQ 359  
Db 301 NSVSLVWVKILVIHDEQKPEVTSNAALTNRNFCWQKHNPSPDRDAEHYDTAILFTRQ 360  
QY 360 DLGSGSOTCDTLGMADVGTCDPSRSCSVIEDGLQAAFTTAHELGHVFNMPHDDAKQAS 419  
Db 361 DLGSGSOTCDTLGMADVGTCDPSRSCSVIEDGLQAAFTTAHELGHVFNMPHDDAKQAS 420  
QY 420 LNGVNQDSHMASMLNLDHSPWSPGSCGYMITSLDNGHGECCLMDKPNQPIQLPGLDLP 479  
Db 421 LNGVNQDSHMASMLNLDHSPWSPGSCGYMITSLDNGHGECCLMDKPNQPIQLPGLDLP 480  
QY 480 TSYDANROCOFTFGSDSKHCPDAASTCTSLMCTGTSGGVLVCQTKHFPWADGTSCEGKW 539  
Db 481 TSYDANROCOFTFGSDSKHCPDAASTCTSLMCTGTSGGVLVCQTKHFPWADGTSCEGKW 540  
QY 540 CINGKCVNKHKHFDTTPHGSWGMWPGDCSRTCSCGGVQYTMRECDNVPVKNCKYCE 599  
Db 541 CINGKCVNKHKHFDTTPHGSWGMWPGDCSRTCSCGGVQYTMRECDNVPVKNCKYCE 600  
QY 600 GKRVRYSCLNEDCDPNNGKTPREBQCEAHNEFSAFSGSPAVIEWIPKYAGVSPKDRCK 659  
Db 601 GKRVRYSCLNEDCDPNNGKTPREBQCEAHNEFSAFSGSPAVIEWIPKYAGVSPKDRCK 660  
QY 660 LICQAKGIGYFVLQPKVVDGTPCSPDSTSCVQGCVKAGCDRIIDSKKPKDKCGVCGG 719  
Db 661 LICQAKGIGYFVLQPKVVDGTPCSPDSTSCVQGCVKAGCDRIIDSKKPKDKCGVCGG 720  
QY 720 NGSTCKKISGVSYSKAPGVHDIITPTGATNIEVKORNGSRNNGSFLAIKAADGTYL 779  
Db 721 NGSTCKKISGVSYSKAPGVHDIITPTGATNIEVKORNGSRNNGSFLAIKAADGTYL 780  
QY 780 NGDYLTLSTLEQDIMYKGVVLRVSGSSAALERSFPLKEPTIQVLTGVNLRPKIKYT 839  
Db 781 NGDYLTLSTLEQDIMYKGVVLRVSGSSAALERSFPLKEPTIQVLTGVNLRPKIKYT 840  
QY 840 YFVKKKESFNAIPFSAWVIEBWEGCSKSCBLGWQRRLVECRDINGQPASCAKEVKPA 899  
Db 841 YFVKKKESFNAIPFSAWVIEBWEGCSKSCBLGWQRRLVECRDINGQPASCAKEVKPA 900  
QY 900 STRPCADHPCPQWOLGENSSCKTCGKYKTSKLASHDGGVLSHDCDPLKPKHFD 959  
Db 901 STRPCADHPCPQWOLGENSSCKTCGKYKTSKLASHDGGVLSHDCDPLKPKHFD 960  
QY 960 FCTMAECS 967  
Db 961 FCTMAECS 968

RESULT 7  
AAW78189  
ID AAW78189 standard; protein; 967 AA.  
XX  
AC AAW78189;  
XX

DT 13-APR-1999 (first entry)

DE Human secreted protein encoded by gene 64 clone HOUQ17.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;  
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
KW developmental abnormality; foetal deficiency; blood; allergy; renal;  
KW immune system; aschma; lymphocytic disease; brain; hepatic; lymphoma;  
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclasts; thymus;  
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

XX Homo sapiens.  
OS  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 40 /label= unknown  
FT Misc-difference 45 /label= unknown  
FT Misc-difference 169 /label= unknown  
FT Misc-difference 293 /label= unknown  
FT Misc-difference 297 /label= unknown  
FT Misc-difference 557 /label= unknown  
XX  
PN WO9856804-A1.  
XX  
PD 17-DEC-1998.  
XX  
PF 11-JUN-1998; 98WO-US012125.  
XX  
PR 13-JUN-1997; 97US-0049547P.  
PR 13-JUN-1997; 97US-0049548P.  
PR 13-JUN-1997; 97US-0049549P.  
PR 13-JUN-1997; 97US-0049550P.  
PR 13-JUN-1997; 97US-0049566P.  
PR 13-JUN-1997; 97US-0049606P.  
PR 13-JUN-1997; 97US-0049607P.  
PR 13-JUN-1997; 97US-0049608P.  
PR 13-JUN-1997; 97US-0049609P.  
PR 13-JUN-1997; 97US-0049610P.  
PR 13-JUN-1997; 97US-0049611P.  
PR 13-JUN-1997; 97US-0050901P.  
PR 13-JUN-1997; 97US-0052989P.  
PR 08-JUL-1997; 97US-0051919P.  
PR 18-AUG-1997; 97US-0055984P.  
PR 12-SEP-1997; 97US-0058668P.  
PR 12-SEP-1997; 97US-0058669P.  
PR 12-SEP-1997; 97US-0058750P.  
PR 12-SEP-1997; 97US-0058971P.  
PR 12-SEP-1997; 97US-0058972P.  
PR 12-SEP-1997; 97US-0058975P.  
PR 02-OCT-1997; 97US-0060834P.  
PR 02-OCT-1997; 97US-0060841P.  
PR 02-OCT-1997; 97US-0060844P.  
PR 02-OCT-1997; 97US-0060865P.  
PR 02-OCT-1997; 97US-0061059P.  
PR 02-OCT-1997; 97US-0061060P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Moore PA, Shi Y, Rosen CA, Ruben SM, Lafleur DW, Olsen HS;  
PI Ebner R, Brewer LA, Young P, Greene JM, Ferrie AM, Yu G, Ni J;  
PI Feng P;  
XX  
DR WPI; 1999-080881/07.  
DR N-PSDB; AAX04374.  
XX  
PT New isolated human genes and the secreted polypeptides they encode -  
PT useful for diagnosis and treatment of e.g. cancers, neurological  
PT disorders, immune diseases, inflammation or blood disorders.  
XX  
PS Claim 11; Page 297-300; 380pp; English.  
XX  
CC This sequence represents a secreted human protein encoded by the gene  
CC clone detailed in the descriptor line. The gene can be used to generate  
CC fusion proteins by linking to the gene to a human immunoglobulin FC  
CC portion (e.g. AAX04302) for increasing the stability of the fused protein  
CC as compared to the human protein only. The invention relates to 86 novel  
CC genes and their fragments (nucleic acid sequences: AAX04311-X04410; amino

CC acid sequences AAW78126-W78225) which are useful for preventing, treating  
CC or ameliorating medical conditions e.g. by protein or gene therapy. Also,  
CC pathological conditions can be diagnosed by determining the amount of the  
CC new polypeptides in a sample or by determining the presence of mutations  
CC in the new polynucleotides. Specific uses are described for each of the  
CC 86 polynucleotides, based on which tissues they are most highly expressed  
CC in (see AAX04311 for described uses)  
XX  
SQ Sequence 967 AA;

Query Match 98.4%; Score 5201; DB 2; Length 967;  
Best Local Similarity 98.3%; Pred. No. 0;  
Matches 951; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 1 MRAVPEGFGRKLGSDMGNAERAPGSRSGFVPFTLLLLAALLAVSDALGRPSDEEEL 60  
DB 1 MRAVPEGFGRKLGSDMGNAERAPGSRSGFVPFTLLLLAALLAVSDALGRPSDEEEL 60  
QY 61 VPELERAPGHGTTLRHLHAFDQDLDPDPSFLAPGFTLQNVGRKSGSDTLPETDL 120  
DB 61 VPELERAPGHGTTLRHLHAFDQDLDPDPSFLAPGFTLQNVGRKSGSETLPETDL 120  
QY 121 AHCFYSGTVNGDPSSAALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAP 180  
DB 121 AHCFYSGTVNGDPSSAALSCEGVRGAFYLLGEAYFIQPLPAASERLXTAAPGEKPPAP 180  
QY 181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETDEDEGTGEDEGPQWSPQDALQGVG 240  
DB 181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETDEDEGTGEDEGPQWSPQDALQGVG 240  
QY 241 OPTGTGSIKRRKRVSSHYRYVETMLVADQSMAPFSGSLKHYLLTLFVVAARLYKHPISRN 300  
DB 241 OPTGTGSIKRRKRVSSHYRYVETMLVADQSMAPFSGSLKHYLLTLFVVAARLYKHPXIRN 300  
QY 301 SVSLVVKILVIHDEQKGPVTSNAALTLRNCWQKOHNPSPDRDAEHYDTAILFTRQD 360  
DB 301 SVSLVVKILVIHDEQKGPVTSNAALTLRNCWQKOHNPSPDRDAEHYDTAILFTRQD 360  
QY 361 LCGSQTCDTLGMAVGTGCDPSRCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420  
DB 361 LCGSQTCDTLGMAVGTGCDPSRCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420  
QY 421 NGVNQDSHMAASLNSLDHSDPSPCSGYMTITSFLDNGHGECCLMDKQNPQIQLPGLDLPGT 480  
DB 421 NGVNQDSHMAASLNSLDHSDPSPCSAYMTITSFLDNGHGECCLMDKQNPQIQLPGLDLPGT 480  
QY 481 SYDANRQCQFTFGEDSKHCPDAASTCTLWCTGTSGGVLCQTKHFPWADGTSCEGKWC 540  
DB 481 SYDANRQCQFTFGEDSKHCPDAASTCTLWCTGTSGGVLCQTKHFPWADGTSCEGKWC 540  
QY 541 INGVKVNKHKHEDTTFHGSWGMWGPWDCSRCTCGGVQVYTMRECDNPVKNKGKCEG 600  
DB 541 INGVKVNKHKHEDTTFHGSWGMWGPWDCSRCTCGGVQVYTMRECDNPVKNKGKCEG 600  
QY 601 KRVRYSCLNEDCDPNNGKTPREQCEAHNEFSAKSGSGPAVEWIPKYAGVSPKDRCKL 660  
DB 601 KRVRYSCLNEDCDPNNGKTPREQCEAHNEFSAKSGSGPAVEWIPKYAGVSPKDRCKL 660  
QY 661 ICQAKGIGYFVLQKVPVVDGTPCSPDSTSVQCGQKAVGCDRIIDSKKPDKCGVCGGN 720  
DB 661 ICQAKGIGYFVLQKVPVVDGTPCSPDSTSVQCGQKAVGCDRIIDSKKPDKCGVCGGN 720  
QY 721 GSTCKK1SGSVTSAPGVHDIITPTGATNLEVKORNGSRNNGSFLAIKAADGTILN 780  
DB 721 GSTCKK1SGSVTSAPGVHDIITPTGATNLEVKORNGSRNNGSFLAIKAADGTILN 780  
QY 781 GDYTLSTLEQD1MYKGVVLYSGSSAALERIRSPSLKEPLTIQVLTGVALRPKIKYTY 840  
DB 781 GDYTLSTLEQD1MYKGVVLYSGSSAALERIRSPSLKEPLTIQVLTGVALRPKIKYTY 840  
QY 841 FVKKKKGSFNAIPTFSAWV1BEWGECSKSCBLGQORLVECRDINGOPASECAKEVPAS 900  
DB 841 FVKKKKGSFNAIPTFSAWV1BEWGECSKSCBLGQORLVECRDINGOPASECAKEVPAS 900

QY 901 TRPCADHPCPQWQLGEWSSCSKTGKGYKKTSLKLSHDSGLSHDSCDPLKKPKHFIDF 960  
DB 901 TRPCADHPCPQWQLGEWSSCSKTGKGYKKTSLKLSHDSGLSHDSCDPLKKPKHFIDF 960  
QY 961 CTMAECS 967  
DB 961 CTMAECS 967

RESULT 8  
ADA57139

ID ADA57139 standard; protein; 967 AA.  
XX ADA57139;  
AC ADA57139;  
XX 20-NOV-2003 (first entry)  
DT 20-NOV-2003 (first entry)  
XX Human secreted protein #422.  
DE Human secreted protein #422.  
XX immunosuppressive; antiinflammatory; antiasthmatic; antiallergic;  
KW cytosolic; cerebroprotective; neuroprotective; nootropic;  
KW cardiovascular; antiarteriosclerotic; gene therapy;  
KW human secreted protein; immune disorder; inflammation;  
KW respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;  
KW inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;  
KW multiple sclerosis; ischaemic brain injury; Parkinson's disease;  
KW Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;  
KW triple helix formation; antisense gene therapy; forensic biology.  
XX Homo sapiens.  
OS Homo sapiens.  
XX WO2002102994-A2.  
PN WO2002102994-A2.  
XX 27-DEC-2002.  
PD 27-DEC-2002.  
XX 19-MAR-2002; 2002WO-US008278.  
PF 19-MAR-2002; 2002WO-US008278.  
XX 21-MAR-2001; 2001US-0277340P.  
PR 21-MAR-2001; 2001US-0277340P.  
PR 19-JUL-2001; 2001US-0306171P.  
PR 13-NOV-2001; 2001US-0331287P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Ruben SM;  
PI Rosen CA, Ruben SM;  
XX WPI; 2003-167512/16.  
DR WPI; 2003-167512/16.  
XX N-PSDB; ADA56243.  
PT New human secreted polypeptides and polynucleotides, useful for  
PT diagnosing, treating or preventing e.g. immune disorders, inflammatory  
PT conditions, respiratory disorders, cancers, CNS disorders, or  
PT neurodegenerative disorders.  
XX Claim 13; SEQ ID NO 1329; 1754pp; English.

XX The invention relates to 592 new human secreted polypeptides useful for  
XX diagnosing, treating or preventing e.g. immune disorders, inflammatory  
XX conditions, respiratory disorders, cancers, CNS disorders, or  
XX neurodegenerative disorders, or polypeptides comprising an amino acid  
XX sequence at least 95% identical to the new sequences. The polypeptides,  
XX antibodies or antibody fragments that bind to the polypeptides, nucleic  
XX acids encoding the polypeptides, agonists or antagonists that binds to  
XX the polypeptide, are useful in preparing diagnostic or pharmaceutical  
XX compositions for diagnosing, treating or preventing an e.g. immune  
XX disorders, inflammatory conditions (e.g. inflammatory bowel disease,  
XX nephritis or Crohn's disease), respiratory disorders (e.g. asthma and  
XX allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders  
XX (e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative  
XX disorders (e.g. Parkinson's disease or Alzheimer's disease), and  
XX cardiovascular disorders (e.g. atherosclerosis or myocarditis). The  
XX polynucleotides are useful for chromosome identification, chromosome  
XX mapping, for controlling gene expression through triple helix formation



or antisense DNA or RNA, in gene therapy, for identifying individuals from minute biological samples, in forensic biology, and as hybridization probes. The polypeptides are useful for as molecular weight markers on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels, to raise antibodies, for testing biological activities, and for treating or preventing neural disorders, immune system disorders, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, proliferative and/or cancerous diseases. This sequence corresponds to one of the polypeptide of the invention. Note: The sequence data for this patent did form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ

Query Match 98.4%; Score 5201; DB 6; Length 967;  
Best Local Similarity 98.3%; Pred. No. 0;  
Matches 951; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 1 MQRVPEGFGRKLGSDMGNAERAPGSRSGFVPVTLALLAALLAVSDALGRPSDEEL 60  
DB 1 MQRVPEGFGRKLGSDMGNAERAPGSRSGFVPVTLALLAALLAVSDALGRPSDEEL 60  
QY 61 VPELERVPGHGTTRLRHLHAFQQLDLDVPPSSFLAPGFTLQVGRKSGSTPLPETDL 120  
DB 61 VPELERVPGHGTTRLRHLHAFQQLDLELRPDSSFLAPGFTLQVGRKSGSTPLPETDL 120  
QY 121 AHCFYSGTVNGDPSSAAALSLCEGVGAPYLLGEAYFTQPLPAASERLATAPGKPPAP 180  
DB 121 AHCFYSGTVNGDPSSAAALSLCEGVGAPYLLGEAYFTQPLPAASERLATAPGKPPAP 180  
QY 181 LQFHLRRNRQGDVGTGCVVDDEPRPTGKATEDEDEGEDEGEQWSPQDPALQGVG 240  
DB 181 LQFHLRRNRQGDVGTGCVVDDEPRPTGKATEDEDEGEDEGEQWSPQDPALQGVG 240  
QY 241 QPTGTGSIKKRFVSSHRYVETMLVADQSMABFHGSLKHYLLTLFVSVAARLYKHPSTRN 300  
DB 241 QPTGTGSIKKRFVSSHRYVETMLVADQSMABFHGSLKHYLLTLFVSVAARLYKHPSTRN 300  
QY 301 SVSLVVKTLVTHDEQKPEVTSNAALTILRNECNKQHNPPSDRDAEHYDTAILFTROD 360  
DB 301 SVSLVVKTLVTHDEQKPEVTSNAALTILRNECNKQHNPPSDRDAEHYDTAILFTROD 360  
QY 361 LCGSOTCDTLGNADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQCASL 420  
DB 361 LCGSOTCDTLGNADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQCASL 420  
QY 421 NGVNQDSHMMASMLNLDHSQWSPSCSYMITSFLDNGHGECLMDKPNQPIQLPGDLPGT 480  
DB 421 NGVNQDSHMMASMLNLDHSQWSPSCSYMITSFLDNGHGECLMDKPNQPIQLPGDLPGT 480  
QY 481 SYDANRQCFTEGDSKPCDPAASTCSTLWCTGTSGGVLVCTQKHPFPAWDTSCGEGKWC 540  
DB 481 SYDANRQCFTEGDSKPCDPAASTCSTLWCTGTSGGVLVCTQKHPFPAWDTSCGEGKWC 540  
QY 541 INGKCNKHKRHFTDPPHSGWMGMPGDCSRTCGGVQVYTMRECDNPVPKNGKCYCEG 600  
DB 541 INGKCNKHKRHFTDPPHSGWMGMPGDCSRTCGGVQVYTMRECDNPVPKNGKCYCEG 600  
QY 601 KRVRYSNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWPKYAGVSPKDRCKL 660  
DB 601 KRVRYSNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWPKYAGVSPKDRCKL 660  
QY 661 ICQAKGIGYFFVLQPKVWDGTCSPDSTSVQVQGVKAGCDRIIDSKKPKGCVCGGN 720  
DB 661 ICQAKGIGYFFVLQPKVWDGTCSPDSTSVQVQGVKAGCDRIIDSKKPKGCVCGGN 720  
QY 721 GSTCKKISGTVSAKPGYHDIITPTGATNIEVKORNGSRNNGSFLAIKAADGTIYLN 780  
DB 721 GSTCKKISGTVSAKPGYHDIITPTGATNIEVKORNGSRNNGSFLAIKAADGTIYLN 780  
QY 781 GDYTLSTLEQDIMYKGVVLRYSGSSAALERTSRFSPLKEPLTIQVLTGVNLRPKIKYTY 840

Db 781 GDYTLSTLEQDIMYKGVVLRYSGSSAALERTSRFSPLKEPLTIQVLTGVNLRPKIKYTY 840  
QY 841 FVKKKESFNAPITFSAWVIEWGECSKSCBLGWQRRLVECRDINGQPASECAKEVKPAS 900  
DB 841 FVKKKESFNAPITFSAWVIEWGECSKSCBLGWQRRLVECRDINGQPASECAKEVKPAS 900  
QY 901 TRPCADHPCPQWQJGEWSSCKTCGKYKTSYLSKLSHGGVLSHSDCDPLKKPKHFIDF 960  
DB 901 TRPCADHPCPQWQJGEWSSCKTCGKYKTSYLSKLSHGGVLSHSDCDPLKKPKHFIDF 960  
QY 961 CTMAECS 967  
DB 961 CTMAECS 967  
RESULT 9  
ADA41003  
ID ADA41003 standard; protein; 967 AA.  
XX  
AC ADA41003;  
DT 20-NOV-2003 (first entry)  
XX  
DE Human secreted protein.  
XX  
KW Human; secreted protein; cancer; hyperproliferative disorder;  
KW rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;  
KW anaemia; allergic reaction; asthma; cardiovascular disorder;  
KW wound healing; cytostatic; immunosuppressive; neuroprotective;  
KW antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory;  
KW vulnery; cardiant; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO2002102993-A2.  
XX  
PD 27-DEC-2002.  
XX  
PF 19-MAR-2002; 2002WO-US008123.  
XX  
PR 21-MAR-2001; 2001US-0277340P.  
PR 19-JUL-2001; 2001US-0306171P.  
PR 13-NOV-2001; 2001US-0331287P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Ruben SM;  
XX  
DR WPI; 2003-175238/17.  
XX  
PT New human secreted proteins and nucleic acid molecules, useful for preparing a diagnostic or pharmaceutical composition for diagnosing, preventing or treating cancer or other hyperproliferative disorder, asthma, allergies or AIDS.  
XX  
PS Claim 1; SEQ ID NO 1385; 3205pp; English.  
XX  
CC The invention relates to novel genes ADA39629-ADA40565 and proteins ADA40566-ADA41501 for human secreted proteins, useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The polypeptides, nucleic acid molecules, antibodies or their fragments, and agonists or antagonists that bind to the polypeptide are useful for preparing a diagnostic or pharmaceutical composition for diagnosing or treating cancer or other hyperproliferative disorder. The polypeptides and nucleic acid molecules are also useful for detecting, preventing, diagnosing, prognosticating, treating or ameliorating cancer or other hyperproliferative disorders including neoplasms, autoimmune disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic anaemia), haematopoietic or haematological disorders (e.g. anaemia, thrombocytopenia), allergic reactions including asthma or eczema, inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory bowel disease or Crohn's disease), neurodegenerative disorders (e.g.

CC Alzheimer's disease or Parkinson's disease), cardiovascular disorders  
 CC (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial,  
 CC fungal or viral infections including HIV/AIDS), or wound healing and  
 CC disorders of epithelial cell proliferation. The nucleic acids are also  
 CC useful for chromosome identification, radiation hybrid mapping or long-  
 CC range restriction mapping, as molecular weight markers, or as  
 CC hybridization or diagnostic probes. The polypeptides and antibodies are  
 CC useful for providing immunological probes for differential identification  
 CC of the tissues immunohistochemistry assays. Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX

SQ Sequence 967 AA;

Query Match 98.4%; Score 5201; DB 6; Length 967;  
 Best Local Similarity 98.3%; Pred. No. 0;  
 Matches 951; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 1 MORAVPEGFRKLGSDMGNAERAPGSRSGVPVPTLLALLAALLAVSDALGRPSDEEL 60  
 DB 1 MORAVPEGFRKLGSDMGNAERAPGSRSGVPVPTLLALLAALLAVSDALGRPSDEEL 60  
 QY 61 VPELERVPGHGTTLRLHLAFDQQLDLDVPPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120  
 DB 61 VPELERAPGHGTTLRLHLAFDQQLDLDVPPDSSFLAPGFTLQNVGRKSGSETPLPETDL 120  
 QY 121 AHCFYSGTVNGDPSAAALSCGVRGAFYLLGRIAYFIQPLPAASERLATAPGEKPPAP 180  
 DB 121 AHCFYSGTVNGDPSAAALSCGVRGAFYLLGRIAYFIQPLPAASERLATAPGEKPPAP 180  
 QY 181 LOFHLLRRNRQDVGCTCGVVDDEPRPTGKAETDEDEGTGEGEPQWSPQDPAALQGVG 240  
 DB 181 LOFHLLRRNRQDVGCTCGVVDDEPRPTGKAETDEDEGTGEGEPQWSPQDPAALQGVG 240  
 QY 241 OPTGTGSRIRKRFVSSHRYVETMLVADQSMAEFFHSGGLKHYLLTLFVSAARLYKHPSTRN 300  
 DB 241 OPTGTGSRIRKRFVSSHRYVETMLVADQSMAEFFHSGGLKHYLLTLFVSAARLYKHPSTRN 300  
 QY 301 SVSLVVKILVHBEQKPEVTSNAALTIRFCNWKQHNPPSDRDAEHYDTAILFTROD 360  
 DB 301 SVSLVVKILVHBEQKPEVTSNAALTIRFCNWKQHNPPSDRDAEHYDTAILFTROD 360  
 QY 361 LCGSQTCTLGMADVTCDSRSCSVLEDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420  
 DB 361 LCGSQTCTLGMADVTCDSRSCSVLEDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420  
 QY 421 NGVNQDSHMMASMLNSLDHSPWSPSCSAYMITSLDNGHGECCLMDKPNPIQLPGDLPGT 480  
 DB 421 NGVNQDSHMMASMLNSLDHSPWSPSCSAYMITSLDNGHGECCLMDKPNPIQLPGDLPGT 480  
 QY 481 SYDANRQOFTFGEDSKHCPDAASTCTSLMCTGTSGGVLVQCTKHPFWDATSGCGKWC 540  
 DB 481 SYDANRQOFTFGEDSKHCPDAASTCTSLMCTGTSGGVLVQCTKHPFWDATSGCGKWC 540  
 QY 541 INGCNVNKHKHPDTPPHGSGWGMGPGDCSRCTCGGQVQVTMRECONVPKNGKCYEG 600  
 DB 541 INGCNVKTDKHPDTPPHGSGWGMGPGDCSRCTCGGQVQVTMRECONVPKNGKCYEG 600  
 QY 601 KRVRYSCLNDCDNNKGTFRQCEAHNFFSKASFGSGPAVEMIPKAGVSPKDRCKL 660  
 DB 601 KRVRYSCLNDCDNNKGTFRQCEAHNFFSKASFGSGPAVEMIPKAGVSPKDRCKL 660  
 QY 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVCCQGVKAGCDRIIDSKKFPDKGVCGGN 720  
 DB 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVCCQGVKAGCDRIIDSKKFPDKGVCGGN 720  
 QY 721 GSTCKKISGVTSAKPGVHDIITPTGATNLEVKORNGSRNNGSFLAIKAADGTILN 780  
 DB 721 GSTCKKISGVTSAKPGVHDIITPTGATNLEVKORNGSRNNGSFLAIKAADGTILN 780  
 QY 781 GDYTLSTLEQDIMYKGVVLRYSGSSAALERIRSFSPLEPTIQLVTGVNALRPKIKYTY 840  
 DB 781 GDYTLSTLEQDIMYKGVVLRYSGSSAALERIRSFSPLEPTIQLVTGVNALRPKIKYTY 840

DB 781 GDYTLSTLEQDIMYKGVVLRYSGSSAALERIRSFSPLEPTIQLVTGVNALRPKIKYTY 840  
 QY 841 FVKKKESFNAIPTFSAWVIEEWGECSSKSELGHQWRRLVECRDINGQPASECAKEVPAS 900  
 DB 841 FVKKKESFNAIPTFSAWVIEEWGECSSKSELGHQWRRLVECRDINGQPASECAKEVPAS 900  
 QY 901 TRPCADHPCPQWOLGEWSSCKTCGKYKTKTSLKLSHDSGLSHDSCDPLKKPKHFIDF 960  
 DB 901 TRPCADHPCPQWOLGEWSSCKTCGKYKTKTSLKLSHDSGLSHDSCDPLKKPKHFIDF 960  
 QY 961 CTMAECS 967  
 DB 961 CTMAECS 967

RESULT 10  
 ADB91631  
 ID ADB91631 standard; protein; 967 AA.

XX AC ADB91631;  
 XX AC ADB91631;  
 DT 04-DEC-2003 (first entry)  
 XX Human secreted protein #SEQ ID 577.  
 DE Secreted protein; gene therapy; antidiabetic; diabetes; human.  
 XX Homo sapiens.  
 XX WO2003004622-A2.  
 XX 16-JAN-2003.  
 XX 19-MAR-2002; 2002WO-US008124.  
 XX 21-MAR-2001; 2001US-0277340P.  
 PR 19-JUL-2001; 2001US-0306171P.  
 PR 13-NOV-2001; 2001US-0331287P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Ruben SM;  
 XX WPI; 2003-229407/22.  
 XX Nucleic acid encoding a human secreted protein is useful in diagnosing or  
 PT treating diabetes or conditions related to diabetes.  
 XX Claim 3; SEQ ID NO 577; 1537pp; English.

XX The invention relates to isolated nucleic acid molecules ADB91065-  
 CC ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-  
 CC ADB91834. Also disclosed is a recombinant vector comprising a  
 CC polynucleotide of the invention, and a recombinant host cell comprising  
 CC the recombinant vector. The polypeptide of the invention is useful in  
 CC identifying a binding partner by contacting the polypeptide with a  
 CC binding partner, and determining whether the binding partner increases or  
 CC decreases activity of the polypeptide. The polypeptide, polynucleotide,  
 CC antibody or its fragment, agonist or antagonist are useful for preparing  
 CC a pharmaceutical composition for diagnosing or treating diabetes or  
 CC conditions related to diabetes. The present sequence is that of the human  
 CC immunoglobulin Fc portion used to generate fusion proteins, increasing  
 CC the stability of the fused protein as compared to the secreted protein  
 CC only. Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX

SQ Sequence 967 AA;

Query Match 98.4%; Score 5201; DB 7; Length 967;  
 Best Local Similarity 98.3%; Pred. No. 0;  
 Matches 951; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

```
QY 1 MORAVPEGFGRRKLGSDMGNAERAPGSRSGFVPVPTLLLLAALLAVSDALGRPSEDEEL 60
Db 1 MORAVPEGFGRRKLGSDMGNAERAPGSRSGFVPVPTLLLLAALLAVSDALGRPSEDEEL 60
QY 61 VPELERVPGHGTTRRLRHAFDQDLDPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
Db 61 VPELERVPGHGTTRRLRHAFDQDLDPDSSFLAPGFTLQNVGRKSGSETLPETDL 120
QY 121 AHCFYSGTVNGDPSSAAALSCEGVRGAFYLLGEAYFTQPLPAASERLATAPGKPPAP 180
Db 121 AHCFYSGTVNGDPSSAAALSCEGVRGAFYLLGEAYFTQPLPAASERLATAPGKPPAP 180
QY 181 LQFHLRRNRQDVGCTCGVDDDEPRPTGKAEDEDESGTEDEGPGQWSPDPAALQGVG 240
Db 181 LQFHLRRNRQDVGCTCGVDDDEPRPTGKAEDEDESGTEDEGPGQWSPDPAALQGVG 240
QY 241 OPTGTGSIKKRFVSSHRVETMLVADQSMAEFHGSLKHYLLTLFSVAARLYKHPSTRN 300
Db 241 OPTGTGSIKKRFVSSHRVETMLVADQSMAEFHGSLKHYLLTLFSVAARLYKHPSTRN 300
QY 301 SVSLVVKILVIHDEQKGPVTSNAALTILRNFQWQKQHNPPSDRDAEHYDTAILFTTRQD 360
Db 301 SVSLVVKILVIHDEQKGPVTSNAALTILRNFQWQKQHNPPSDRDAEHYDTAILFTTRQD 360
QY 361 LCGSQTCDTLGADVGTVCDSRSRCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQKQASL 420
Db 361 LCGSQTCDTLGADVGTVCDSRSRCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQKQASL 420
QY 421 NGVNQDSHMMASMLNLDHSQWSPSCSYMYITSFLDNGHGCECLMDKQNPITQLPGDLPGT 480
Db 421 NGVNQDSHMMASMLNLDHSQWSPSCSYMYITSFLDNGHGCECLMDKQNPITQLPGDLPGT 480
QY 481 SYDANRQCOFTFGEDSKCPDAASTCTLWCTGTSGGVLVCOQTHFPWADGTSCGEGKWC 540
Db 481 SYDANRQCOFTFGEDSKCPDAASTCTLWCTGTSGGVLVCOQTHFPWADGTSCGEGKWC 540
QY 541 INGKCVNKHRRKHFDTFPHGSGMWGMPGDCSRTCGGVQVYTMRECDNPVPKNGKCYCEG 600
Db 541 INGKCVXKTDKHFDTFPHGSGMWGMPGDCSRTCGGVQVYTMRECDNPVPKNGKCYCEG 600
QY 601 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660
Db 601 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660
QY 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVQGVQCVKAGCDRIIDSKKKFKDCGVCNGN 720
Db 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVQGVQCVKAGCDRIIDSKKKFKDCGVCNGN 720
QY 721 GSTCKKISGSVTSAPGYHDIITPTGATNIEVKQNRQGRNNGSFLAIKAADGTIYL 780
Db 721 GSTCKKISGSVTSAPGYHDIITPTGATNIEVKQNRQGRNNGSFLAIKAADGTIYL 780
QY 781 GDTLTSTLEQDIIMYGVLRYSGSAALERISFSPLEPLTIQVLTGNALRPKIKTY 840
Db 781 GDTLTSTLEQDIIMYGVLRYSGSAALERISFSPLEPLTIQVLTGNALRPKIKTY 840
QY 841 FVKKKESFNALPTFSAWVIBEWGCSKSCELGWQRRLVECRDINGQPASECACKEVKPAS 900
Db 841 FVKKKESFNALPTFSAWVIBEWGCSKSCELGWQRRLVECRDINGQPASECACKEVKPAS 900
QY 901 TRPCADHPCPQWOLGEWSCKTCGKYKTKSLKCLSHDGGVLSHSDCDPLKPKPHFIDF 960
Db 901 TRPCADHPCPQWOLGEWSCKTCGKYKTKSLKCLSHDGGVLSHSDCDPLKPKPHFIDF 960
QY 961 CTMAECS 967
Db 961 CTMAECS 967
```

RESULT 11

ADC74267

ID ADC74267 standard; protein; 967 AA.

XX

```
AC ADC74267;
XX 01-JAN-2004 (first entry)
XX Human secreted protein - SEQ ID 900.
XX antianemic; antirheumatic; antiarthritic; antiinflammatory; antithyroid;
XX antidabetic; immunosuppressive; dermatologic; nephrotropic;
XX antiparkinsonian; neuroprotective; nootropic; antibacterial; virucide;
XX fungicide; antiparasitic; antiarteriosclerotic; vulnerary; cytostatic;
XX haemopoietic; haematologic; anaemia; autoimmune disorder;
XX rheumatoid arthritis; inflammation; Grave's disease; diabetes;
XX systemic lupus erythematosus; glomerulonephritis; neurodegenerative;
XX Parkinson's; Alzheimer's; wound; hyperproliferative; atherosclerosis;
XX cancer; bacterial; viral; fungal; parasitic infection; gene therapy;
XX human.
XX Homo sapiens.
OS WO2003038063-A2.
XX 08-MAY-2003.
XX 19-MAR-2002; 2002WO-US008277.
XX 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM;
XX WPI: 2003-430516/40.
XX N-PSDB; ADC73652.
XX New human secreted polypeptide for diagnosing, preventing or treating
XX hemopoietic or hematologic disorders (e.g. anemia), autoimmune
XX disorders (e.g. diabetes) or hyperproliferative disorders (e.g. cancer or
XX atherosclerosis).
XX Claim 16; SEQ ID NO 900; 2272pp; English.
XX The invention relates to a novel human secreted polypeptide comprising a
XX defined sequence given in the specification. The polypeptide, nucleic
XX acid molecule, antibody, agonist or antagonist of the invention may be
XX useful for preparing a composition for diagnosing or treating a
XX haemopoietic or haematologic disorder such as anaemia, autoimmune
XX disorders such as rheumatoid arthritis, inflammation, Grave's disease,
XX diabetes, systemic lupus erythematosus or glomerulonephritis,
XX neurodegenerative disorders including Parkinson's disease and Alzheimer's
XX disease, wounds and hyperproliferative disorders including
XX atherosclerosis or cancer, as well as bacterial, viral, fungal or
XX parasitic infections. The polypeptide may also be used during gene
XX therapy procedures and for identifying a binding partner by contacting
XX the polypeptide with a binding partner and determining whether the
XX binding partner increases or decreases the activity of the polypeptide.
XX The current sequence is that of the human secreted protein of the
XX invention.
XX SQ Sequence 967 AA;
Query Match 98.4%; Score 5201; DB 7; Length 967;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 951; Conservative 4; Mismatches 12; Indels 0; Gaps 0;
QY 1 MORAVPEGFGRRKLGSDMGNAERAPGSRSGFVPVPTLLLLAALLAVSDALGRPSEDEEL 60
Db 1 MORAVPEGFGRRKLGSDMGNAERAPGSRSGFVPVPTLLLLAALLAVSDALGRPSEDEEL 60
QY 61 VPELERVPGHGTTRRLRHAFDQDLDPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
Db 61 VPELERVPGHGTTRRLRHAFDQDLDPDSSFLAPGFTLQNVGRKSGSETLPETDL 120
```

QY	121	AHCFYSGTVNGDPSSAALSICEGVGAFYLLGEAYFIQPLPAASERLATAAPCEKPPAP	180
DB	121	AHCFYSGTVNGDPSSAALSICEGVGAFYLLGEAYFIQPLPAASERLATAAPCEKPPAP	180
QY	181	LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETEDDEGTEDEGPGQWSPQDPALQGVG	240
DB	181	LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETEDDEGTEDEGPGQWSPQDPALQGVG	240
QY	241	QPTGTGSIRKKRFVSSHRYVETMLVADQSMAEFHGSLKHYLLTLFSAARLYKHPISRN	300
DB	241	QPTGTGSIRKKRFVSSHRYVETMLVADQSMAEFHGSLKHYLLTLFSAARLYKHPISRN	300
QY	301	SVSLVVKILVIHDEQKPEVTSNAALTIRFCNWKQHNPSDRDAEHYDTAILFTROD	360
DB	301	SVSLVVKILVIHDEQKPEVTSNAALTIRFCNWKQHNPSDRDAEHYDTAILFTROD	360
QY	361	LCGSQTCDTLGMADVTCVDPSSRSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQCASL	420
DB	361	LCGSQTCDTLGMADVTCVDPSSRSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQCASL	420
QY	421	NGVNQDSHMAASMLSNLDHSPFSPGSGYMITSLFDNGHGCLMDKPNQPIQLPGDLPGT	480
DB	421	NGVNQDSHMAASMLSNLDHSPFSPGSGYMITSLFDNGHGCLMDKPNQPIQLPGDLPGT	480
QY	481	SYDANRQCOFTFGEDSKHCPDAASTCSLTWCTGTSGVLVCTQKHPWADGTSCEGKWC	540
DB	481	SYDANRQCOFTFGEDSKHCPDAASTCSLTWCTGTSGVLVCTQKHPWADGTSCEGKWC	540
QY*	541	INGKCVNKNRKHFDPTPHGSGWGMWPGMDCSRTCGGQVQYTMRECONVPKNGKCEG	600
DB	541	INGKCVNKNRKHFDPTPHGSGWGMWPGMDCSRTCGGQVQYTMRECONVPKNGKCEG	600
QY	601	KRVYRSCNLEDCPDNNGKTPREQCEAHNBFSAKASFGSGPAVEWIPKYAGVSPKDRCKL	660
DB	601	KRVYRSCNLEDCPDNNGKTPREQCEAHNBFSAKASFGSGPAVEWIPKYAGVSPKDRCKL	660
QY	661	ICQAKGIGYFVLQPKVVDGTPCSPDSTSCVQGCYKAGCDRIIDSKKFDKGCVCGN	720
DB	661	ICQAKGIGYFVLQPKVVDGTPCSPDSTSCVQGCYKAGCDRIIDSKKFDKGCVCGN	720
QY	721	GSTCKKISGVSATKPGVHDIITPTGATNIEVQRNQRGSRNNGSFLAIKAADGTILN	780
DB	721	GSTCKKISGVSATKPGVHDIITPTGATNIEVQRNQRGSRNNGSFLAIKAADGTILN	780
QY	781	GDYTLSTLEQIMYKGVVLRYSGSAALERTRSFSLKEPLTIQVLTGVGNALRPKIYTY	840
DB	781	GDYTLSTLEQIMYKGVVLRYSGSAALERTRSFSLKEPLTIQVLTGVGNALRPKIYTY	840
QY	841	FVKKKESFNAIPTFSAWVIEWGECSKSCELGWORRLVECRDINGOPASECAKEVPAS	900
DB	841	FVKKKESFNAIPTFSAWVIEWGECSKSCELGWORRLVECRDINGOPASECAKEVPAS	900
QY	901	TRPCADHPCPQWQLGEWSSCKTCGKYKKTSLKCLSHDGVLSHDSCLDKPKPHFIDF	960
DB	901	TRPCADHPCPQWQLGEWSSCKTCGKYKKTSLKCLSHDGVLSHDSCLDKPKPHFIDF	960
QY	961	CTMAECS 967	
DB	961	CTMAECS 967	
RESULT 12			
ADD37948			
ID	ADD37948	standard; protein; 967 AA.	
XX	AC	ADD37948;	
XX	XX		
XX	DT	15-JAN-2004 (first entry)	
XX	DE	Human secreted protein #131.	
XX	XX	human secreted protein; Antiallergic; Antiinflammatory; Antibacterial;	
KW			

KW	Anti-HIV; Cytostatic; Immunosuppressive; Hemostatic.
XX	
OS	Homo sapiens.
XX	
PN	WO200290526-A2.
XX	
PD	14-NOV-2002.
XX	
PF	19-MAR-2002; 2002WO-US008279.
XX	
PR	21-MAR-2001; 2001US-0277340P.
PR	19-JUL-2001; 2001US-0306171P.
PR	13-NOV-2001; 2001US-0331287P.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Rosen CA, Ruben SM;
XX	
DR	WPI; 2003-140218/13.
XX	
PT	New human secreted proteins and nucleic acid molecules, useful for preparing a diagnostic or pharmaceutical composition for diagnosing or treating allergic or asthmatic disorders, or related immediate hypersensitivity disorders.
PT	
PT	
XX	
PS	Claim 1; SEQ ID NO 430; 1323pp; English.
XX	
CC	The present invention relates to an isolated polypeptide or human secreted protein. The polypeptides, nucleic acid molecules, antibodies or their fragments, and agonists or antagonists that bind are useful for preparing a diagnostic or pharmaceutical composition for diagnosing or treating allergic or asthmatic disorders. The polypeptide is also useful for identifying a binding partner by contacting the polypeptide with a binding partner, and determining whether the binding partner increases or decreases the activity of the polypeptide. The polypeptides and nucleic acid molecules are also useful for detecting, preventing, diagnosing, prognosticating, treating or ameliorating inflammatory disorders. Neoplastic diseases, wound healing and disorders of epithelial cell proliferation, immune disorders, cardiovascular disorders, blood-related disorders, infectious diseases, endocrine disorders, or gastrointestinal disorders. The nucleic acids are also useful for chromosome identification, radiation hybrid mapping or long-range restriction mapping, as molecular weight markers, or as hybridization or diagnostic probes. The polypeptides and antibodies are useful for providing immunological probes for differential identification of the tissues immunohistochemistry assays. The present sequence represents a human secreted protein.
XX	
SQ	Sequence 967 AA;
	Query Match 98.4%; Score 5201; DB 7; Length 967;
	Best Local Similarity 98.3%; Pred. No. 0;
	Matches 951; Conservative 4; Mismatches 12; Indels 0; Gaps 0;
QY	1 MQRAVPEGFORRKLGSMDGNAERAPGSRSPGVPPTLLLLAAALLAVSDALGRPSEDEEL 60
DB	1 MQRAVPEGFORRKLGSMDGNAERAPGSRSPGVPPTLLLLAAALLAVSDALGRPSEDEEL 60
QY	61 VPELERVPGHGTTRLLRLHAFDQDLDDVPDSSFLAPGFTLQNVGRKSGDTPLPETDL 120
DB	61 VPELERVPGHGTTRLLRLHAFDQDLDDVPDSSFLAPGFTLQNVGRKSGDTPLPETDL 120
QY	121 AHCFYSGTVNGDPSSAALSICEGVGAFYLLGEAYFIQPLPAASERLATAAPCEKPPAP 180
DB	121 AHCFYSGTVNGDPSSAALSICEGVGAFYLLGEAYFIQPLPAASERLATAAPCEKPPAP 180
QY	181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETEDDEGTEDEGPGQWSPQDPALQGVG 240
DB	181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETEDDEGTEDEGPGQWSPQDPALQGVG 240
QY	241 QPTGTGSIRKKRFVSSHRYVETMLVADQSMAEFHGSLKHYLLTLFSAARLYKHPISRN 300
DB	241 QPTGTGSIRKKRFVSSHRYVETMLVADQSMAEFHGSLKHYLLTLFSAARLYKHPISRN 300

QY 301 SVSLVVVKILVTHDEQKGEVTSNAALTLRNFQCNWQKHNPSPDRDAEHYDTAILFTROD 360  
Db 301 SVSLVVVKILVTHDEQKGEVTSNAALTLRNFQCNWQKHNPSPDRDAEHYDTAILFTROD 360  
QY 361 LCGSQTCDTLGWADVGTVCDPSPSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACSL 420  
Db 361 LCGSQTCDTLGWADVGTVCDPSPSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACSL 420  
QY 421 NGVNQDSHMAWSMLNLDHSHQSPSCSGYMTSFLDNGHGECLMDKPNPQIOLPGDLPGT 480  
Db 421 NGVNQDSHMAWSMLNLDHSHQSPSCSGYMTSFLDNGHGECLMDKPNPQIOLPGDLPGT 480  
QY 481 SYDANRQCFTFGEDSKHCPDAASTCSTLWCTGTSGVLVCOQTKHPFWADGTCSEGRWC 540  
Db 481 SYDANRQCFTFGEDSKHCPDAASTCSTLWCTGTSGVLVCOQTKHPFWADGTCSEGRWC 540  
QY 541 INKCVNKHRRKHFDPFFHSGWGMWGPWGDGSRCTCGGVGYTMRCDNVPKNGKCYCG 600  
Db 541 INKCVNKHRRKHFDPFFHSGWGMWGPWGDGSRCTCGGVGYTMRCDNVPKNGKCYCG 600  
QY 601 KRVRYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660  
Db 601 KRVRYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660  
QY 661 ICOAKGIGYFFVLQPKVVDGTPCSPDSTSVQVQCVKAGCDRIIDSKKKFKDCKGVCGN 720  
Db 661 ICOAKGIGYFFVLQPKVVDGTPCSPDSTSVQVQCVKAGCDRIIDSKKKFKDCKGVCGN 720  
QY 721 GSTCKKISGVSATSAKGYHDIITIPGATNIEVKQNRQSRNNGSFLAIIKAADGTIILN 780  
Db 721 GSTCKKISGVSATSAKGYHDIITIPGATNIEVKQNRQSRNNGSFLAIIKAADGTIILN 780  
QY 781 GYTLSTLEQDMYKGVLYRSGSSAALIRISFSPLEPLTIQVLTGVALRPKIKYTY 840  
Db 781 GYTLSTLEQDMYKGVLYRSGSSAALIRISFSPLEPLTIQVLTGVALRPKIKYTY 840  
QY 841 FVKKKESFNAIPTFSAMVIEBWGECSKSELGQRRRLVECRDINGQPASECAKEVKPAS 900  
Db 841 FVKKKESFNAIPTFSAMVIEBWGECSKSELGQRRRLVECRDINGQPASECAKEVKPAS 900  
QY 901 TRPCADHPCPQOLGBWSSCSKTCGKYKTSIKLSHDGGVLSHDSDPLKKPKHFIDF 960  
Db 901 TRPCADHPCPQOLGBWSSCSKTCGKYKTSIKLSHDGGVLSHDSDPLKKPKHFIDF 960  
QY 961 CTMAECS 967  
Db 961 CTMAECS 967  
RESULT 13  
AAV49501  
ID AAY49501 standard; protein; 950 AA.  
XX  
AC AAY49501;  
XX  
DT 10-JAN-2000 (first entry)  
XX  
DE Human METH1 protein.  
XX  
KW Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;  
KW cancer; diagnosis; hyperproliferative disorder; autoimmune disease;  
KW angiogenesis inhibitor; abnormal wound healing; inflammation;  
KW rheumatoid arthritis; psoriasis; endometrial bleeding disorder;  
KW diabetic retinopathy; macula degeneration; haemangioma; detection;  
KW arterial-venous malformation; immune deficiency.  
XX  
OS Homo sapiens.  
XX  
PN WO9937660-A1.  
XX  
PD 29-JUL-1999.  
XX

PF 22-JAN-1999; 99WO-US001313.  
XX  
PR 23-JAN-1998; 98US-0072298P.  
PR 28-AUG-1998; 98US-0098539P.  
XX  
PA (IRUE/) IRUELA-ARISPE L.  
PA (HAST/) HASTINGS G A.  
PA (RUBE/) RUBEN S M.  
XX  
PI Iruela-Arispe L, Hastings GA, Ruben SM;  
XX  
XX WPI: 1999-590684/50.  
DR N-PSDB; AA232000.  
XX  
PT New isolated metalloprotease thrombospondin polypeptides, useful for  
PT treating hyperproliferative disorders, cancers or autoimmune disorders.  
XX  
PS Claim 10; Fig 1; 457pp; English.  
XX  
CC AA232000 and AA232001 encode, and AAY49501 and AAY49502 represent, human  
CC metalloprotease thrombospondin (METH) proteins METH1 and METH2  
CC respectively. METH1 and METH2 have been found to be potent inhibitors of  
CC angiogenesis both in vitro and in vivo. They can be used for treating  
CC cancer and other disorders related to angiogenesis including abnormal  
CC wound healing, inflammation, rheumatoid arthritis, psoriasis, endometrial  
CC bleeding disorders, diabetic retinopathy, some forms of macula  
CC degeneration, haemangiomas, and arterial-venous malformations. They may  
CC be useful in treating deficiencies or disorders of the immune system, by  
CC activating or inhibiting the proliferation, differentiation, or  
CC mobilisation (chemotaxis) of immune cells. The etiology of these immune  
CC deficiencies or disorders may be genetic, somatic, such as cancer or some  
CC autoimmune disorders, acquired (e.g. by chemotherapy or toxins), or  
CC infectious. They can also be used to treat inflammatory conditions, both  
CC chronic and acute conditions. The products can also be used for detection  
CC and diagnosis. AA232002 to AA232080, and AAY49503 to AAY49511 represent  
CC sequences given in the exemplification of the present invention  
XX  
SQ Sequence 950 AA;  
Query Match 97.3%; Score 5144; DB 2; Length 950;  
Best Local Similarity 98.9%; Pred No. 0;  
Matches 940; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
QY 18 MGNAPRPGSRSGFGVPVPTLLLAALLAVSDALGRPSSEDEELVPELVRPGHGTTLR 77  
Db 1 MGNAPRPGSRSGFGVPVPTLLLAALLAVSDALGRPSSEDEELVPELVRPGHGTTLR 60  
QY 78 LHAFOQLDLVPPDSSFLAFGFTLQNVGRKSGSTPLPETDLAHCFSGTNGDPSSAA 137  
Db 61 LHAFOQLDLVPPDSSFLAFGFTLQNVGRKSGSTPLPETDLAHCFSGTNGDPSSAA 120  
QY 138 ALSLCEGVGRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLOFHLRRNRQGDVGT 197  
Db 121 ALSLCEGVGRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLOFHLRRNRQGDVGT 180  
QY 198 CGVDDPRPTGKAFTEDGEGEPQWSPDPAALQGVQPTGTGSRKRFVSSH 257  
Db 181 CGVDDPRPTGKAFTEDGEGEPQWSPDPAALQGVQPTGTGSRKRFVSSH 240  
QY 258 RYVETMLVADQSMABFHGSLKHYLLTLFSAARLYKHPSIRNSVSLVVKILVTHDEQK 317  
Db 241 RYVETMLVADQSMABFHGSLKHYLLTLFSAARLYKHPSIRNSVSLVVKILVTHDEQK 300  
QY 318 GPEVTSNAALTLRNFQCNWQKHNPSPDRDAEHYDTAILFTQDLCSGTCDTLGMADVGT 377  
Db 301 GPEVTSNAALTLRNFQCNWQKHNPSPDRDAEHYDTAILFTQDLCSGTCDTLGMADVGT 360  
QY 378 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACSLNGVNDSHMWSMLSNL 437  
Db 361 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACSLNGVNDSHMWSMLSNL 420  
QY 438 DHSQWPSCSGYMTSFLDNGHGECLMDKPNPQIOLPGDLFGTSDANRQCFTFGESDK 497  
XX

Db 421 DHSQWSPSCAYMITSFLDNGHGECLMDKPNQIOLPGDLPGTSYDANRQCQFTFGEDSK 480  
 QY 498 HCPDAASTCTLWCTGTSGGVLVVCOTKHPPWADGTSCEGKWCINGKCVNKNHRKHFDTP 557  
 Db 481 HCPDAASTCTLWCTGTSGGVLVVCOTKHPPWADGTSCEGKWCINGKCVNKNHRKHFDTP 540  
 QY 558 FHGSGMGPWGDGCSRTCGGGVQYTMRECDNVPKNGGKYCEGKRVYRSCNLEDCPDNN 617  
 Db 541 FHGSGMGPWGDGCSRTCGGGVQYTMRECDNVPKNGGKYCEGKRVYRSCNLEDCPDNN 600  
 QY 618 GKTFRQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLIICQAKGIGYFFVLQPKV 677  
 Db 601 GKTFRQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLIICQAKGIGYFFVLQPKV 660  
 QY 678 VDGTPCSPDSTVVCVQGCYKAGCDRIIDSCKKFKDKCGVCGNGSTCKKISGSVTSAPKG 737  
 Db 661 VDGTPCSPDSTVVCVQGCYKAGCDRIIDSCKKFKDKCGVCGNGSTCKKISGSVTSAPKG 720  
 QY 738 YHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYILNGDYTLSTLEQDIDMYKGV 797  
 Db 721 YHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYILNGDYTLSTLEQDIDMYKGV 780  
 QY 798 VLRYSGSSAALERIRSFSPKLEPLTIQVLTGVALRPKI KYTFYVKKKESFNAIPTFSA 857  
 Db 781 VLRYSGSSAALERIRSFSPKLEPLTIQVLTGVALRPKI KYTFYVKKKESFNAIPTFSA 840  
 QY 858 WVEEWGECSCSELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHPCPQWLGEW 917  
 Db 841 WVEEWGECSCSELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHPCPQWLGEW 900  
 QY 918 SSCSKTCGKYKTSKLSKSHDGGVLSHDSCDPLKPKHFDICTMAECS 967  
 Db 901 SSCSKTCGKYKTSKLSKSHDGGVLSHDSCDPLKPKHFDICTMAECS 950

RESULT 14  
 AAB73549  
 ID AAB73549 standard; protein; 950 AA.  
 XX AC AAB73549;  
 XX DT 07-AUG-2001 (first entry)  
 XX DE Human ADAM-type metalloprotease MDTs4, SEQ ID NO:4.  
 XX KW Human; MDTs4; ADAM-type metalloprotease; drug screening;  
 XX OS A Disintegrin And Metalloprotease; cancer; arthritis.  
 XX OS Homo sapiens.  
 XX PN JP2001017183-A.  
 XX PD 23-JAN-2001.  
 XX PF 09-JUL-1999; 99JP-00196584.  
 XX PR 09-JUL-1999; 99JP-00196584.  
 XX PA (YAMA ) YAMANOUCHI PHARM CO LTD.  
 XX WPI; 2001-275950/29.  
 XX DR N-PSDB; AAB20224.  
 XX PT A new metal protease and its preparation for use as an anti-cancer and  
 XX PT anti-arthritis therapeutic.  
 XX PS Claim 1; Page 12-14; 22pp; Japanese.  
 XX CC The invention relates to the novel human ADAM (A Disintegrin And  
 CC Metalloprotease)-type metalloproteases MDTs4 (AAB73549) and MDTs5  
 CC (AAB73550). The metalloproteases can be used for the treatment of cancers  
 CC and arthritis. The invention also relates to the genes encoding MDTs4 and  
 CC MDTs5, vectors and host cells containing the MDTs4 or MDTs5 genes, the

CC recombinant production of MDTs4 and MDTs5, and antibody specific for  
 CC MDTs4 or MDTs5, and methods of screening for compounds which modulate the  
 CC activity of MDTs4 and/or MDTs5. The present sequence represents human  
 CC MDTs4  
 XX SQ Sequence 950 AA;  
 Query Match 97.3%; Score 5144; DB 4; Length 950;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 940; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
 QY 18 MGNERAPGSRSGFQVPTLLILAAALLAVSDALGRPSEDEELVPELERVPHGHTFLR 77  
 Db 1 MGNERAPGSRSGFQVPTLLILAAALLAVSDALGRPSEDEELVPELERVPHGHTFLR 60  
 QY 78 LHAFDQQLDLVPPDSSFLAPGFTLQNVGRKSGSDTFLPETDLAHCYFSGTVNGDPSAA 137  
 Db 61 LHAFDQQLDLVPPDSSFLAPGFTLQNVGRKSGSETPLPETDLAHCYFSGTVNGDPSAA 120  
 QY 138 ALSCEGVRGAFYLLGEAYFIQPLPAASERLATAARKEKPPAPLOFHLLRNROGDYGGT 197  
 Db 121 ALSCEGVRGAFYLLGEAYFIQPLPAASERLATAARKEKPPAPLOFHLLRNROGDYGGT 180  
 QY 198 CGVVDDEPRPTGKAETDEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGE 257  
 Db 181 CGVVDDEPRPTGKAETDEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGE 240  
 QY 258 RYVETMLVADQSMAEFHGSLKHYLLTLFVSAARLYKHPISIRNSVLSLVKILVHDEQK 317  
 Db 241 RYVETMLVADQSMAEFHGSLKHYLLTLFVSAARLYKHPISIRNSVLSLVKILVHDEQK 300  
 QY 318 GPEVTSNAALTNRNFCNWKOHNPSPDRDAEHYDTAILFTRODLCSGOTCDTLGMADVGT 377  
 Db 301 GPEVTSNAALTNRNFCNWKOHNPSPDRDAEHYDTAILFTRODLCSGOTCDTLGMADVGT 360  
 QY 378 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACASLNGVQDSHMASLSNL 437  
 Db 361 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACASLNGVQDSHMASLSNL 420  
 QY 438 DHSQWSPSCGYMITSFIDNGHGECLMDKPNQIOLPGDLPGTSYDANRQCQFTFGEDSK 497  
 Db 421 DHSQWSPSCGYMITSFIDNGHGECLMDKPNQIOLPGDLPGTSYDANRQCQFTFGEDSK 480  
 QY 498 HCPDAASTCTLWCTGTSGGVLVVCOTKHPPWADGTSCEGKWCINGKCVNKNHRKHFDTP 557  
 Db 481 HCPDAASTCTLWCTGTSGGVLVVCOTKHPPWADGTSCEGKWCINGKCVNKNHRKHFDTP 540  
 QY 558 FHGSGMGPWGDGCSRTCGGGVQYTMRECDNVPKNGGKYCEGKRVYRSCNLEDCPDNN 617  
 Db 541 FHGSGMGPWGDGCSRTCGGGVQYTMRECDNVPKNGGKYCEGKRVYRSCNLEDCPDNN 600  
 QY 618 GKTFRQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLIICQAKGIGYFFVLQPKV 677  
 Db 601 GKTFRQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLIICQAKGIGYFFVLQPKV 660  
 QY 678 VDGTPCSPDSTVVCVQGCYKAGCDRIIDSCKKFKDKCGVCGNGSTCKKISGSVTSAPKG 737  
 Db 661 VDGTPCSPDSTVVCVQGCYKAGCDRIIDSCKKFKDKCGVCGNGSTCKKISGSVTSAPKG 720  
 QY 738 YHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYILNGDYTLSTLEQDIDMYKGV 797  
 Db 721 YHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYILNGDYTLSTLEQDIDMYKGV 780  
 QY 798 VLRYSGSSAALERIRSFSPKLEPLTIQVLTGVALRPKI KYTFYVKKKESFNAIPTFSA 857  
 Db 781 VLRYSGSSAALERIRSFSPKLEPLTIQVLTGVALRPKI KYTFYVKKKESFNAIPTFSA 840  
 QY 858 WVEEWGECSCSELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHPCPQWLGEW 917  
 Db 841 WVEEWGECSCSELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHPCPQWLGEW 900  
 QY 918 SSCSKTCGKYKTSKLSKSHDGGVLSHDSCDPLKPKHFDICTMAECS 967  
 Db 901 SSCSKTCGKYKTSKLSKSHDGGVLSHDSCDPLKPKHFDICTMAECS 950

Db 901 SSCSKTCGKYKRSLSKLSHDSGVLSHESCDPLKKPKHFIDFCTMAECS 950

RESULT 15

AAB50002  
ID AAB50002 standard; protein; 950 AA.

XX AC AAB50002;

XX DT 19-MAR-2001 (first entry)

XX DE Human METH1.

XX KW Human; METH1; metalloprotease; thrombospondin; angiogenesis inhibition;  
KW cancer therapy; benign tumour; ocular angiogenic disease;  
KW rheumatoid arthritis; psoriasis; wound healing; endometriosis;  
KW vasculogenesis; granulation; hypertrophic scar; nonunion fracture;  
KW scleroderma, trachoma; vascular adhesion; myocardial angiogenesis;  
KW coronary collateral; cerebral collateral; arteriovenous malformation;  
KW ischaemic limb angiogenesis; Osler-Webber syndrome;  
KW plaque neovascularisation; telangiectasia; haemophilic joint;  
KW angiofibroma; fibromuscular dysplasia; wound granulation;  
KW Crohn's disease; atherosclerosis; birth control.

XX OS Homo sapiens.

XX PN WO200071577-A1.

XX PD 30-NOV-2000.

XX PF 25-MAY-2000; 2000WO-US014462.

XX PR 25-MAY-1999; 99US-00318208.

XX PR 20-JUL-1999; 99US-0144882P.

XX PR 10-AUG-1999; 99US-0147823P.

XX PR 13-AUG-1999; 99US-00372658.

XX PR 22-DEC-1999; 99US-0171503P.

XX PR 22-FEB-2000; 2000US-0183792P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX PA (SMUK-) SMITHKLINE BEECHAM CORP.

XX PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX PA (IRUE/) IRUELA-ARISPE L.

XX PA (HAST-) HASTINGS G A.

XX PA (RUBE/) RUBEN S M.

XX PA (JONA/) JONAK Z L.

XX PA (TRUL/) TRULLI S H.

XX PA (FORN/) FORNWALD J A.

XX PA (TERR/) TERRETT J A.

XX PI Iruela-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH;

XX PI Fornwald JA, Terrett JA;

XX DR WPI; 2001-025136/03.

XX DR N-PSDB; AAC90057.

XX METH1 and METH2 polynucleotides and encoded peptides, used to inhibit  
XX angiogenesis in the treatment of disorders such as cancer, rheumatoid  
XX arthritis and psoriasis.

XX Claim 15; Fig 1; 769pp; English.

XX The present sequence is human METH1 (ME for metalloprotease and TH for  
XX thrombospondin). METH1 can be used for inhibiting angiogenesis in an  
XX individual, and for treating cancer, benign tumours, an ocular angiogenic  
XX disease, rheumatoid arthritis, psoriasis, delayed wound healing,  
XX endometriosis, vasculogenesis, granulations, hypertrophic scars, nonunion  
XX fractures, scleroderma, trachoma, vascular adhesions, myocardial  
XX angiogenesis, coronary collaterals, cerebral collaterals, arteriovenous  
XX malformations, ischaemic limb angiogenesis, Osler-Webber syndrome, plaque  
XX neovascularisation, telangiectasia, haemophilic joints, angiofibroma,  
XX fibromuscular dysplasia, wound granulation, Crohn's disease or  
XX atherosclerosis. METH1 can also be used in birth control. METH1 can also

CC be used in diagnostic methods for the prognosis of cancer

XX

SQ Sequence 950 AA;

Query Match 97.3%; Score 5144; DB 4; Length 950;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 940; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 18 MGNAERAPGSRSGFVPPTLLALLAALAVSDALGRPSEDEBELVVPPELVRVFGHGTTLR 77

DB 1 MGNAERAPGSRSGFVPPTLLALLAALAVSDALGRPSEDEBELVVPPELVRVFGHGTTLR 60

QY 78 LHAFOQDLDPDPSSFLAFQFTLQNVGRKSGSDTLPETDLAHCFSVTGNDGPPSAA 137

DB 61 LHAFOQDLLELRDPSSFLAFQFTLQNVGRKSGSETPLPETDLAHCFSVTGNDGPPSAA 120

QY 138 ALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPQFHLRLNRQGDVG 197

DB 121 ALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPQFHLRLNRQGDVG 180

QY 198 CGVVDDEPRPTGKAETDEDETEGEDEGPQSPQDPALQGVGPTGTGSIKKRFVSSH 257

DB 181 CGVVDDEPRPTGKAETDEDETEGEDEGPQSPQDPALQGVGPTGTGSIKKRFVSSH 240

QY 258 RYVETMLVADQSMARFHSGLKHYLLTLFSVAARLYKHPISIRNSVSLVVKILVIHDSK 317

DB 241 RYVETMLVADQSMARFHSGLKHYLLTLFSVAARLYKHPISIRNSVSLVVKILVIHDSK 300

QY 318 GPEVTSNAALTNRNFCWQKQHNPPSDRDAHYDTAILFTQDLGCGSOTCDTLGNADVGT 377

DB 301 GPEVTSNAALTNRNFCWQKQHNPPSDRDAHYDTAILFTQDLGCGSOTCDTLGNADVGT 360

QY 378 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQKASLNGVNDOSHMSLSNL 437

DB 361 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQKASLNGVNDOSHMSLSNL 420

QY 438 DHSQPSWSPSCGYMITSFLDNHGCELMKPNQPTQLPGDLPGTSDANRQCQFTFGEDSK 497

DB 421 DHSQPSWSPSCGYMITSFLDNHGCELMKPNQPTQLPGDLPGTSDANRQCQFTFGEDSK 480

QY 498 HCPDAASTCSTLWCTGTGGVLCQTKHFPWADGTCGEGKWCINGKCNKNHRKHDP 557

DB 481 HCPDAASTCSTLWCTGTGGVLCQTKHFPWADGTCGEGKWCINGKCNKNHRKHDP 540

QY 558 FHSGWGMWGPWGDCSRTCGGVQYTMRECDNPVPKNGKCYCEGKRVYRSCNLEDCPDNN 617

DB 541 FHSGWGMWGPWGDCSRTCGGVQYTMRECDNPVPKNGKCYCEGKRVYRSCNLEDCPDNN 600

QY 618 GKTFRQCEAHNBFSSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKV 677

DB 601 GKTFRQCEAHNBFSSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKV 660

QY 678 VDGTPCSPDSTVCQGVKAGCDRIIDSKKFKDCGCGNGSTCKKISGSVTSKPG 737

DB 661 VDGTPCSPDSTVCQGVKAGCDRIIDSKKFKDCGCGNGSTCKKISGSVTSKPG 720

QY 738 YHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTIILNGDYTLSTLEQDQIMYKGV 797

DB 721 YHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTIILNGDYTLSTLEQDQIMYKGV 780

QY 798 VLRYSGSSAALERTSRSPLEPTIQLVTGNALRPKIYTYFVKKKESFNAIPTFSA 857

DB 781 VLRYSGSSAALERTSRSPLEPTIQLVTGNALRPKIYTYFVKKKESFNAIPTFSA 840

QY 858 WVIIEWGECSSCELGWQRRLVECRDINGQPASCAKEVKPASTRCPADHPCPQWQJGEW 917

DB 841 WVIIEWGECSSCELGWQRRLVECRDINGQPASCAKEVKPASTRCPADHPCPQWQJGEW 900

QY 918 SSCSKTCGKYKRSLSKLSHDSGVLSHESCDPLKKPKHFIDFCTMAECS 967

DB 901 SSCSKTCGKYKRSLSKLSHDSGVLSHESCDPLKKPKHFIDFCTMAECS 950

Search completed: August 22, 2005, 10:15:39  
Job time : 176 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2005, 10:09:59 ; Search time 44 Seconds  
(without alignments)  
1640.583 Million cell updates/sec

Title: US-09-989-687-126

Perfect score: 5287  
Sequence: 1 MORAVPEGFRKLGSDMGN.....CDPLKKPKHFIDFCTMAECS 967

Scoring table: BLOSUMP2  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/1aa/5A-COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B-COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/6A-COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/6B-COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5232	99.0	967	4	US-09-130-491-2
2	5140	97.2	949	4	US-09-568-559-2
3	4260.5	80.6	950	4	US-09-321-987B-4
4	3922	74.2	727	4	US-09-445-023A-1
5	3656	69.2	727	4	US-09-445-023A-12
6	3297	62.4	608	4	US-09-130-491-13
7	2794	52.8	551	4	US-09-130-491-16
8	2471.5	46.7	950	4	US-10-009-332-1
9	2265.5	42.9	905	3	US-09-369-364A-9
10	2124	40.2	837	4	US-09-122-126B-2
11	2124	40.2	837	4	US-09-634-286A-2
12	2124	40.2	837	4	US-10-247-685-2
13	1924	36.4	930	4	US-09-122-126B-15
14	1924	36.4	930	4	US-09-634-286A-15
15	1924	36.4	930	4	US-10-247-685-15
16	1909.5	36.1	930	3	US-09-369-364A-2
17	1731.5	32.8	1882	3	US-09-369-364A-13
18	1608.5	30.4	2150	4	US-09-321-987B-2
19	1605.5	30.4	874	3	US-09-369-364A-15
20	1598	30.2	2165	4	US-09-800-729-155
21	1390.5	26.3	481	3	US-09-130-491-8
22	1293	26.3	518	3	US-09-369-364A-22
23	1287	24.3	1104	4	US-09-981-953A-4
24	1278.5	24.2	1224	4	US-09-930-872-4
25	1278.5	24.2	1224	4	US-10-217-774-4
26	1214.5	23.0	908	4	US-09-963-791-2
27	1207.5	22.8	997	3	US-09-369-364A-7

28	1176	22.2	1081	3	US-09-369-364A-17	Sequence 17, Appl
29	1151.5	21.8	757	4	US-09-963-791-24	Sequence 24, Appl
30	1100.5	20.8	1211	4	US-09-949-016-11401	Sequence 11401, A
31	1096	20.7	969	4	US-09-321-987B-5	Sequence 5, Appl
32	1083.5	20.5	1211	4	US-09-491-522-5	Sequence 5, Appl
33	1059	20.0	1205	4	US-09-491-522-11	Sequence 11, Appl
34	1029.5	19.5	859	3	US-09-369-364A-5	Sequence 5, Appl
35	1017	19.2	770	4	US-09-981-953A-2	Sequence 2, Appl
36	1001	18.9	1039	4	US-09-949-016-7859	Sequence 7859, Ap
37	781	14.8	589	4	US-09-963-791-12	Sequence 12, Appl
38	712.5	13.6	438	4	US-09-963-791-22	Sequence 22, Appl
39	712.5	13.5	245	3	US-09-369-364A-11	Sequence 11, Appl
40	595	11.3	1745	4	US-09-800-729-89	Sequence 89, Appl
41	593.5	11.2	507	4	US-09-963-791-10	Sequence 10, Appl
42	568.5	10.8	525	3	US-09-369-364A-21	Sequence 21, Appl
43	566.5	10.7	724	4	US-09-784-358-8	Sequence 8, Appl
44	566.5	10.7	845	4	US-09-784-358-12	Sequence 12, Appl
45	566.5	10.7	1691	4	US-09-784-358-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-09-130-491-2  
; Sequence 2, Application US/09130491  
; Patent No. 6416974  
; GENERAL INFORMATION:  
; APPLICANT: Holtzman, Douglas A.  
; APPLICANT: Goodearl, Andrew D.J.  
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83  
; FILE REFERENCE: 09404/041001  
; CURRENT APPLICATION NUMBER: US/09/130,491  
; CURRENT FILING DATE: 1998-08-07  
; EARLIER APPLICATION NUMBER: US 60/058,108  
; EARLIER FILING DATE: 1997-09-05  
; EARLIER APPLICATION NUMBER: US 60/054,961  
; EARLIER FILING DATE: 1997-08-06  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 967  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-130-491-2

Query Match	99.0%;	Score 5232;	DB 4;	Length 967;
Best Local Similarity	99.0%;	Fred. No. 0;		
Matches 957;	Conservative 4;	Mismatches 6;	Indels 0;	Gaps 0;
Qy	1	MORAVPEGFRKLGSDMGNAPGSRSGFVPPTLLLLAALLAVSDALGRPGEDEEL	60	
Db	1	MORAVPEGFRKLGSDMGNAPGSRSGFVPPTLLLLAALLAVSDALGRPGEDEEL	60	
Qy	61	VPELERYPGHGTTLRLHAFDQQLDLVDPDSFAPGFTLQNVGRKSGSDTLPETDL	120	
Db	61	VPELERYPGHGTTLRLHAFDQQLDLVDPDSFAPGFTLQNVGRKSGSETLPETDL	120	
Qy	121	AHCFYSGTVNGDPSAAALSLCEGVRGAFYLLGAYFTQPLPAASERLATAAPGKPPAP	180	
Db	121	AHCFYSGTVNGDPSAAALSLCEGVRGAFYLLGAYFTQPLPAASERLATAAPGKPPAP	180	
Qy	181	LQFHLLRNROGDVGTCGVVDDPRPTGKAETDEDESGTEGEGEGPQWSPQDPALQVG	240	
Db	181	LQFHLLRNROGDVGTCGVVDDPRPTGKAETDEDESGTEGEGEGPQWSPQDPALQVG	240	
Qy	241	QPTGTGSRKRRFVSSHRYVETMLVADQSMAEFHGSLKHYLLTLFVSAARLYKHPISRN	300	
Db	241	QPTGTGSRKRRFVSSHRYVETMLVADQSMAEFHGSLKHYLLTLFVSAARLYKHPISRN	300	
Qy	301	SVSLVVKILVIHDEQKGPEVTSNAALTLRNFCWQKOHNPSPDRDAEHYOTAILFTQD	360	
Db	301	SVSLVVKILVIHDEQKGPEVTSNAALTLRNFCWQKOHNPSPDRDAEHYOTAILFTQD	360	

QY 361 LCGSOTCDTLGMADYGVCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACASL 420  
 Db 361 LCGSOTCDTLGMADYGVCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACASL 420  
 QY 421 NGVNQDSHMMASMLNLDHSHQSPWSCSGYMITSFLLDNGHGECLMDKPNQPIQLPGDLPGT 480  
 Db 421 NGVNQDSHMMASMLNLDHSHQSPWSCSGYMITSFLLDNGHGECLMDKPNQPIQLPGDLPGT 480  
 QY 481 SYDANROCOFTFGEDSKHCPCDAASTCTLWCTGTSGGVLVQTKHFPWADGTSCEGKWC 540  
 Db 481 SYDANROCOFTFGEDSKHCPCDAASTCTLWCTGTSGGVLVQTKHFPWADGTSCEGKWC 540  
 QY 541 INKCKVKNHRKHFDTPFHSGWGMWPGDCSRTCGGVQVYTMRECDNPVKNKGKCEG 600  
 Db 541 INKCKVKNHRKHFDTPFHSGWGMWPGDCSRTCGGVQVYTMRECDNPVKNKGKCEG 600  
 QY 601 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660  
 Db 601 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660  
 QY 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVQGVQCVKAGCDRIIDSKKKFDKCGVCVCGN 720  
 Db 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVQGVQCVKAGCDRIIDSKKKFDKCGVCVCGN 720  
 QY 721 GSTCKKISGSVTSAPKGYVHDIIITPTGATNIEVKORNORSGRNNGSFLAIKAAADGTYILN 780  
 Db 721 GSTCKKISGSVTSAPKGYVHDIIITPTGATNIEVKORNORSGRNNGSFLAIKAAADGTYILN 780  
 QY 781 GDTLSTLEQDIIIMYKGVVLRYSGSSAALERSFPLKEPLTIQVLTGNALRPKIKYTY 840  
 Db 781 GDTLSTLEQDIIIMYKGVVLRYSGSSAALERSFPLKEPLTIQVLTGNALRPKIKYTY 840  
 QY 841 FVKKKESFNAIPTSAWVIEBWGCSKSCBLGWORRLVECRDINGQPASECAKEVKNPAS 900  
 Db 841 FVKKKESFNAIPTSAWVIEBWGCSKSCBLGWORRLVECRDINGQPASECAKEVKNPAS 900  
 QY 901 TRPCADHPCPQWOLGEWSSCKTCGKYKTSKCLSHDGGVLSHSCDPLKKPKHFD 960  
 Db 901 TRPCADHPCPQWOLGEWSSCKTCGKYKTSKCLSHDGGVLSHSCDPLKKPKHFD 960  
 QY 961 CTMAECS 967  
 Db 961 CTMAECS 967

RESULT 2  
 US-09-568-559-2  
 ; Sequence 2, Application US/09568559  
 ; Patent No. 6649377  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Klonowski, Paul  
 ; APPLICANT: Allard, John  
 ; APPLICANT: Heller, Renu  
 ; APPLICANT: Van Wart, Harold  
 ; TITLE OF INVENTION: Human Aggrecanase and Nucleic Acid  
 ; TITLE OF INVENTION: Compositions Encoding the Same  
 ; FILE REFERENCE: ROCH-002  
 ; CURRENT APPLICATION NUMBER: US/09/568,559  
 ; CURRENT FILING DATE: 2000-05-09  
 ; PRIOR APPLICATION NUMBER: 60/133,343  
 ; PRIOR FILING DATE: 1999-05-10  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 949  
 ; TYPE: PRT  
 ; ORGANISM: human  
 US-09-568-559-2

Query Match 97.2%; Score 5140; DB 4; Length 949;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 939; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 19 GNAERAPCSRSFGVPTILLILAAALLAVSDALGRPSBEDELAVPELVRVPGHGTTRRL 78  
 Db 1 GNAERAPCSRSFGVPTILLILAAALLAVSDALGRPSBEDELAVPELVRVPGHGTTRRL 60  
 QY 79 HAFQOQLDLVDPDSSFLAPGFTLQNVGRKSGSTPLPETDLAHCYFSGTVNGPSSAAA 138  
 Db 61 HAFQOQLDLVDPDSSFLAPGFTLQNVGRKSGSTPLPETDLAHCYFSGTVNGPSSAAA 120  
 QY 139 LSLCEGVRGAPYLLGEAYFIQPLPAASERLATAAPGKPPAPLOFHLRLRRNQDVGTC 198  
 Db 121 LSLCEGVRGAPYLLGEAYFIQPLPAASERLATAAPGKPPAPLOFHLRLRRNQDVGTC 180  
 QY 199 GVWDEPRPTKATEDEDETEDEGEDEGAQMSPODPALQGVQGTGTSIRKKKPFVSHR 258  
 Db 181 GVWDEPRPTKATEDEDETEDEGEDEGAQMSPODPALQGVQGTGTSIRKKKPFVSHR 240  
 QY 259 YVETMLVADQSMABFHGSGLKHYLLTLFSVAARLYKHPSIRNSVSLVVKILVHDEQK 318  
 Db 241 YVETMLVADQSMABFHGSGLKHYLLTLFSVAARLYKHPSIRNSVSLVVKILVHDEQK 300  
 QY 319 PEVTSNAALTIRNFCNWKOHNPSPDRDAEHYDTAILFTRQDLCSQTCDTLGMADVGT 378  
 Db 301 PEVTSNAALTIRNFCNWKOHNPSPDRDAEHYDTAILFTRQDLCSQTCDTLGMADVGT 360  
 QY 379 CDPSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACASLNGVNQDSHMMASMLNLD 438  
 Db 361 CDPSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACASLNGVNQDSHMMASMLNLD 420  
 QY 439 HSQPSWPCSGYMITSFLLDNGHGECLMDKPNQPIQLPGDLPGTSYDANROCOFTFGEDSKH 498  
 Db 421 HSQPSWPCSGYMITSFLLDNGHGECLMDKPNQPIQLPGDLPGTSYDANROCOFTFGEDSKH 480  
 QY 499 CPDAASTCTLWCTGTSGGVLVQTKHFPWADGTSCEGKWCINGKCVKNHRKHFDTPP 558  
 Db 481 CPDAASTCTLWCTGTSGGVLVQTKHFPWADGTSCEGKWCINGKCVKNHRKHFDTPP 540  
 QY 559 HSGWGMWPGDCSRTCGGVQVYTMRECDNPVKNKGKCEGKRVYRSCNLEDCPDNNG 618  
 Db 541 HSGWGMWPGDCSRTCGGVQVYTMRECDNPVKNKGKCEGKRVYRSCNLEDCPDNNG 600  
 QY 619 KTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKV 678  
 Db 601 KTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKV 660  
 QY 679 DGTPCSPDSTSVQGVQCVKAGCDRIIDSKKKFDKCGVCVCGNGSTCKKISGSVTSAPGY 738  
 Db 661 DGTPCSPDSTSVQGVQCVKAGCDRIIDSKKKFDKCGVCVCGNGSTCKKISGSVTSAPGY 720  
 QY 739 HDIITPTGATNIEVKORNORSGRNNGSFLAIKAAADGTYILNGDYTLSTLEQDIIIMYKGV 798  
 Db 721 HDIITPTGATNIEVKORNORSGRNNGSFLAIKAAADGTYILNGDYTLSTLEQDIIIMYKGV 780  
 QY 799 LRYSGSSAALERSFPLKEPLTIQVLTGNALRPKIKYTYFVKKKESFNAIPTSAW 858  
 Db 781 LRYSGSSAALERSFPLKEPLTIQVLTGNALRPKIKYTYFVKKKESFNAIPTSAW 840  
 QY 859 VIEBWGCSKSCBLGWORRLVECRDINGQPASECAKEVKNPASRCPADHPCPQWOLGEWS 918  
 Db 841 VIEBWGCSKSCBLGWORRLVECRDINGQPASECAKEVKNPASRCPADHPCPQWOLGEWS 900  
 QY 919 SCSKTCGKYKTSKCLSHDGGVLSHSCDPLKKPKHFDICTMAECS 967  
 Db 901 SCSKTCGKYKTSKCLSHDGGVLSHSCDPLKKPKHFDICTMAECS 949

RESULT 3  
 US-09-321-987B-4  
 ; Sequence 4, Application US/09321987B  
 ; Patent No. 6730820  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kimble, Judith E  
 ; APPLICANT: Bielloch, Robert H

; TITLE OF INVENTION: Agent and Method for Modulating Cell Migration

; FILE REFERENCE: 960296.95386  
; CURRENT APPLICATION NUMBER: US/09/321,987B  
; CURRENT FILING DATE: 1999-05-28  
; PRIOR APPLICATION NUMBER: 60/087,170  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: 60/129,023  
; PRIOR FILING DATE: 1999-04-13  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 950  
; TYPE: PRT  
; ORGANISM: Murine  
US-09-321-987B-4

Query Match 80.6%; Score 4260.5; DB 4; Length 950;  
Best Local Similarity 81.6%; Pred. No. 0;  
Matches 781; Conservative 56; Mismatches 105; Indels 15; Gaps 5;

QY 18 MGNARAPGRSGFGVPVPTLLLLAAA---LLAVSDALGRPSEDEELVPELERVFGHG-- 72  
DB 1 MGDVQARARSGLSAHMLLLLSITMLLCARGAHRTEDEELVPLSERAFGHDS 60

QY 73 TTRLRLHAPDQOLDLDVPPDSFLAPGFTLQNVGRKSGSDTLP--PETDLAHCFTYSGTVN 130  
DB 61 TTRLRLDAFGQQLHLKLOPDSGLFAPGFTLQVGRSPGSEAQLDPTGDLAHCFTYSGTVN 120

QY 131 GDPSSAAALSCEGVRGAFYLLGEAYFTOLPL- AASERLATAAPGEKPPAPLOFHLLARN 189  
DB 121 GDPGSAALSCEGVRGAFYLLQGEFFIQAPGVATERLAPVPEESSARPFQFHILRRR 180

QY 190 ROGVDGTCGVVDDEPRPTGKAETDEDEBTEGEDEGPQSDPALQVGVOPTGTGSR 249  
DB 181 RRGSGAKGVNDDELPT- -----SDSRPSQNRNQMPVRDPTPDAGKSPGSGSR 233

QY 250 KKRFSVSHRYVETMLVADQSMAEFHGSLKHVLLTLFSVAARLYKHPSIRNSVSLVVVKI 309  
DB 234 KKRFSVSPRYVETMLVADQSMADFHGSLKHVLLTLFSVAARFYKHPSIRNSISLWVKI 293

QY 310 LVIHDEQKGPVETSNAAALTLRNFQWQKOHNPSPDRDAEHYDTAILFTRQDLCSGQTC 369  
DB 294 LVIYEQKGPVETSNAAALTLRNFQWQKOHNPSPDRDPEHYDTAILFTRQDLCSGHTC 353

QY 370 LGMADVGTCDPSRCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACASLNGVNDSHM 429  
DB 354 LGMADVGTCDPSRCSVIEDDGLQAAFTTAHELGHVFNPHDDAKXGASLNGVTDSHL 413

QY 430 MASMLSNLDHSPQWSPSCGYMITSFLDNGHGCLMDKPNQPIQLPGDLPGTSDANRQC 489  
DB 414 MASMLSSLDHSPQWSPSCSAYMITSFLDNGHGCLMDKPNQPIKLPDLPGLTYDANRQC 473

QY 490 FTFGEDSKHCPDAASTCTLWCTGSGGVLCVQTKHFPWADGTSCEGKWCINGKVCVN 549  
DB 474 FTFGESKHCPCDAASTCTLWCTGSGGVLCVQTKHFPWADGTSCEGKWCYSGKVCNKT 533

QY 550 HKRKHFTDPFHGSGMWGPMGDCSRTC CGGVQVYTMRECDNPVPKNGGKYCEGKRVYRSCN 609  
DB 534 DMKHFTDPFHGSGMWGPMGDCSRTC CGGVQVYTMRECDNPVPKNGGKYCEGKRVYRSCN 593

QY 610 LEDCPDNNKGTTFREEQCEAHNFEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIGY 669  
DB 594 IEDCPDNNKGTTFREEQCEAHNFEFSKASFGNEPTVEWTPKYAGVSPKDRCKLTCEAKGIGY 653

QY 670 FVLQPKVVDGTPCSPDSTSVCVQGCYKAGCDRIIDSKKKPKCGVCGNGSTCKKISG 729  
DB 654 FVLQPKVVDGTPCSPDSTSVCVQGCYKAGCDRIIDSKKKPKCGVCGNGSTCKKMSG 713

QY 730 SVTSKAPGVHDIITPTGATNTEVKORNGRNNGSFLAIIKAAADQTYILNGDYTLSTILE 789  
DB 714 IVTSTRPGVHDIITPTGATNTEVKHNRNGRNNGSFLAIRAADQTYILNGDYTLSTILE 773

QY 790 QDIMYKGVVLYSGSAAALERIRSFPLKEPLTIQVLTGVNALRPKIKYTYFVKKKXSPF 849

RESULT 4

US-09-445-023A-1  
; Sequence 1, Application US/09445023A  
; Patent No. 6565858

; GENERAL INFORMATION:  
; APPLICANT: Hiroee, Kunitaka  
; APPLICANT: Inoguchi, Eiichi  
; APPLICANT: Hakozaaki, Michinori  
; APPLICANT: Ishioaka, Keiko  
; APPLICANT: Ishida, Yukako  
; APPLICANT: Matsushima, Kouji  
; APPLICANT: Kuno, Kouji  
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical  
; FILE OF INVENTION: Composition and method of immunologically analyzing human ADAMTS-1  
; FILE REFERENCE: Q57092  
; CURRENT APPLICATION NUMBER: US/09/445,023A  
; CURRENT FILING DATE: 1999-12-03  
; PRIOR APPLICATION NUMBER: JP 9-160422  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 727  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-445-023A-1

Query Match 74.2%; Score 3922; DB 4; Length 727;  
Best Local Similarity 98.2%; Pred. No. 0;  
Matches 707; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 248 IRKRFVSSHRYVETMLVADQSMAEFHGSLKHVLLTLFSVAARLYKHPSIRNSVSLVVV 307  
DB 8 LRKRFVSSPRYVETMLVADQSMAEFHGSLKHVLLTLFSVAARLYKHPSIRNSVSLVVV 67

QY 308 KILVIHDEQKGPVETSNAAALTLRNFQWQKOHNPSPDRDAEHYDTAILFTRQDLCSGQTC 367  
DB 68 KILVIHDEQKGPVETSNAAALTLRNFQWQKOHNPSPDRDAEHYDTAILFTRQDLCSGQTC 127

QY 368 DTLMADVGTCDPSRCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACASLNGVNDSD 427  
DB 128 DTLMADVGTCDPSRCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACASLNGVNDSD 187

QY 428 HMASMLSNLDHSPQWSPSCGYMITSFLDNGHGCLMDKPNQPIQLPGDLPGTSDANRQ 487  
DB 188 HMASMLSNLDHSPQWSPSCSAYMITSFLDNGHGCLMDKPNQPIQLPGDLPGTLYDANRQ 247

QY 488 CQFTFGEDSKHCPDAASTCTLWCTGSGGVLCVQTKHFPWADGTSCEGKWCINGKVCVN 547  
DB 248 CQFTFGEDSKHCPDAASTCTLWCTGSGGVLCVQTKHFPWADGTSCEGKWCINGKVCVN 307

QY 548 KNHRKHFTDPFHGSGMWGPMGDCSRTC CGGVQVYTMRECDNPVPKNGGKYCEGKRVYRSCN 607  
DB 308 KTRKHFTDPFHGSGMWGPMGDCSRTC CGGVQVYTMRECDNPVPKNGGKYCEGKRVYRSCN 367

QY 608 CNLEDCPDNNKGTTFREEQCEAHNFEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGI 667  
DB 368 CNLEDCPDNNKGTTFREEQCEAHNFEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGI 427

QY 668 GYFFVLQPKVVDGTPCSPDSTSVCVQGCYKAGCDRIIDSKKKPKCGVCGNGSTCKKI 727

```
Db 428 GYFFVLQPKVVDGTPCSPDSTSVCGQCCKAGCDRIIDSKKKFKDKCGVCGGNGSTCKKI 487
Qy 728 SGSVTSAPGYHDIITPTGATNIEVKQNRGSRNNGSFLAIKAADGTYILNGDYTLST 787
Db 488 SGSVTSAPGYHDIITPTGATNIEVKQNRGSRNNGSFLAIKAADGTYILNGDYTLST 547
Qy 788 LEQDIMYKGVILRYSGSSAALRIIRSFSPKLEPLTIQVLTGVNALRPKIKYTYFYVKKKE 847
Db 548 LEQDIMYKGVILRYSGSSAALRIIRSFSPKLEPLTIQVLTGVNALRPKIKYTYFYVKKKE 607
Qy 848 SFNAIPTSASVIEEWGECSKCELGWRRLVECRDINGOPASECAKEVKPASTRPCADH 907
Db 608 SFNAIPTSASVIEEWGECSKCELGWRRLVECRDINGOPASECAKEVKPASTRPCADH 667
Qy 908 PCPQOLGEWSSCKTCGKYKTSIKCLSHDGGVLSHSDCDPLKKPKHFIDFCTMAECS 967
Db 668 PCPQOLGEWSSCKTCGKYKTSIKCLSHDGGVLSHSDCDPLKKPKHFIDFCTMAECS 727

RESULT 5
US-09-445-023A-12
; Sequence 12, Application US/09445023A
; Patent No. 6565858
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiji
; APPLICANT: Hakozaaki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; FILE REFERENCE: Q57092
; CURRENT APPLICATION NUMBER: US/09/445,023A
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-445-023A-12

Query Match 69.2%; Score 3656; DB 4; Length 727;
Best Local Similarity 89.0%; Pred. No. 1.9e-311;
Matches 647; Conservative 42; Mismatches 39; Indels 0; Gaps 0;

Qy 241 OPTGTGSRKRRFVSSHRYVETMLVADQSMAEFHGSLKHYLLTLFSVAARLYKHPISRN 300
Db 1 EPSGPGSRKRRFVSSHRYVETMLVADQSMADFHGSLKHYLLTLFSVAARLYKHPISRN 60
Qy 301 SVSLVVKVILVIHDEQKPEVTSNAALTLRNFCNWKQKHNPPSDRDAEHYDTAILFTROD 360
Db 61 SISLVVVKVILVIHDEQKPEVTSNAALTLRNFCNWKQKHNPPSDRDAEHYDTAILFTROD 120
Qy 361 LCGSQTCDTLGMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACASL 420
Db 121 LCGSHTCDTLGMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKHCA 180
Qy 421 NGVNQDSHMAASMLSLNDHSPWSPCSGYMTITSFLDNGHGECCLMDKPNPIQLPGDLP 480
Db 181 NGVSGDSHMAASMLSLNDHSPWSPCSGYMTITSFLDNGHGECCLMDKPNPIQLPGDLP 240
Qy 481 SYDANRQCQFTFGEDSKHCPDAASTCTSLWCTGTSGGVLCVQTKHFPWADGTSCEGKWC 540
Db 241 LYDANRQCQFTFGEDSKHCPDAASTCTSLWCTGTSGGVLCVQTKHFPWADGTSCEGKWC 300
Qy 541 INKCVNKHKHFDTTPHGSWGMWGPWGDGSRCTCGGVQVYTMRECDNPPKNGKCYCEG 600
Db 301 VSGKCVNKHKHFDTTPHGSWGMWGPWGDGSRCTCGGVQVYTMRECDNPPKNGKCYCEG 360
```

```
Qy 601 KRVRYSNILEDCPDNNKKTFRBQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660
Db 361 KRVRYSNILEDCPDNNKKTFRBQCEAHNEFSKASFGNEPTWTPKYAGVSPDRCKL 420
Qy 661 ICOAKGIGYFFVLQPKVVDGTPCSPDSTSVCGQCCKAGCDRIIDSKKKFKDKCGVCGG 720
Db 421 TCEAKGIGYFFVLQPKVVDGTPCSPDSTSVCGQCCKAGCDRIIDSKKKFKDKCGVCGG 480
Qy 721 GSTCKKISGSVTSAPGYHDIITPTGATNIEVKQNRGSRNNGSFLAIKAADGTYILN 780
Db 481 GSTCKKISGSVTSAPGYHDIITPTGATNIEVKQNRGSRNNGSFLAIKAADGTYILN 540
Qy 781 GDYTLSTLEQDIMYKGVILRYSGSSAALRIIRSFSPKLEPLTIQVLTGVNALRPKIKY 840
Db 541 GNFTLSTLEQDIMYKGVILRYSGSSAALRIIRSFSPKLEPLTIQVLMVGHALRPKIKY 600
Qy 841 FVKKKESFNAIPTSASVIEEWGECSKCELGWRRLVECRDINGOPASECAKEVKPAS 900
Db 601 FVKKKESFNAIPTSASVIEEWGECSKCELGWRRLVECRDINGOPASECAKEVKPAS 660
Qy 901 TRPCADHPCPQOLGEWSSCKTCGKYKTSIKCLSHDGGVLSHSDCDPLKKPKHFIDF 960
Db 661 TRPCADHPCPQOLGEWSSCKTCGKYKTSIKCLSHDGGVLSHSDCDPLKKPKHFIDF 720
Qy 961 CTMAECS 967
Db 721 CTLTQCS 727

RESULT 6
US-09-130-491-13
; Sequence 13, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-130-491-13

Query Match 62.4%; Score 3297; DB 4; Length 608;
Best Local Similarity 90.9%; Pred. No. 4.3e-280;
Matches 599; Conservative 1; Mismatches 5; Indels 54; Gaps 1;
```

```
Qy 309 ILVIHDEQKPEVTSNAALTLRNFCNWKQKHNPPSDRDAEHYDTAILFTRODLCGSOTCD 368
Db 4 ILVIHDEQKPEVTSNAALTLRNFCNWKQKHNPPSDRDAEHYDTAILFTRODLCGSOTCD 63
Qy 369 TLGMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACASLNGVNQDSH 428
Db 64 TLGMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACASLNGVNQDSH 123
Qy 429 MMASMLSLNDHSPWSPCSGYMTITSFLDNGHGECCLMDKPNPIQLPGDLPDGTSYDANRQC 488
Db 124 MMASMLSLNDHSPWSPCSGYMTITSFLDNGHGECCLMDKPNPIQLPGDLPDGTSYDANRQC 183
Qy 489 QFTFGEDSKHCPDAASTCTSLWCTGTSGGVLCVQTKHFPWADGTSCEGKWCINGKCVN 548
Db 184 QFTFGEDSKHCPDAASTCTSLWCTGTSGGVLCVQTKHFPWADGTSCEGKWCINGKCVN 243
```

Qy	549	NHRKHFDTPFHGSGWMGPGWDCSRTCTGGGVQVYTWRECDNPPVXKNGGKYCEGKRVRYSR	608
Db	244	TDRKHFDTPFHGSGWMGPGWDCSRTCTGGGVQVYTWRECDNPPVXKNGGKYCEGKRVRYSR	303
Qy	609	NLEDCPDNNGKTFREOCEAHNEFSAKSFSGSPAWEWIPKYAGVSPKDRCKLICQAKGIG	668
Db	304	NLEDCPDNNGKTFREOCEAHNEFSAKSFSGSPAWEWIPKYAGVSPKDRCKLICQAKGIG	363
Qy	669	YFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSXKXFPDKGVCVGGNGSTCKKIS	728
Db	364	YFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSXKXFPDKGVCVGGNGSTCKKIS	423
Qy	729	GSVTSAKPGYHDIITPTGATNIEVKQRNQRSGRNNGSFLAIKAADGTYILNGDYTLSTL	788
Db	424	GSVTSAKPGYHDIITPTGATNIEVKQRNQRSGRNNGSFLAIKAADGTYILNGDYTLSTL	483
Qy	789	EQDIMKVGVLRYSGSSAALERIRSFSPKLEPITTVQVLTVGNALRPKIXYTVFVKKKES	848
Db	484	EQDIMKVGVLRYSGSSAALERIRSFSPKLEPITTVQVLTVGNALRPKIXYTVFVKKKES	543
Qy	849	FNAIPTSAAWVIEEWGECSSKCELGWQRRLVECRDINGOPASECAKEVKAPESTRPCADHP	908
Db	544	FNAIPTSAAWVIEEWGECSSKCELGWQRRLVECRDINGOPASECAKEVKAPESTRPCADHP	560
Qy	909	CPQWOLGWSWSSCKTCKGKYKXTSLKCLSHDGGVLSHSDCDPLKKPKGPHIDFCTWAECS	967
Db	561	-----CSKTCGKYKKRSLKCLSHDGGVLSHSDCDPLKKPKGPHIDFCTWAECS	608
RESULT 7			
US-09-130-491-16			
; Sequence 16, Application US/09130491			
; Patent No. 6416974			
; GENERAL INFORMATION:			
; APPLICANT: Holtzman, Douglas A.			
; APPLICANT: Goodearl, Andrew D.J.			
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83			
; FILE REFERENCE: 09404/041001			
; CURRENT APPLICATION NUMBER: US/09/130,491			
; CURRENT FILING DATE: 1998-08-07			
; EARLIER APPLICATION NUMBER: US 60/058,108			
; EARLIER FILING DATE: 1997-09-05			
; EARLIER APPLICATION NUMBER: US 60/054,961			
; EARLIER FILING DATE: 1997-08-06			
; NUMBER OF SEQ ID NOS: 16			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 16			
; LENGTH: 551			
; TYPE: PRT			
; ORGANISM: Rattus rattus			
US-09-130-491-16			

Query Match	52.8%	Score	2794;	DB	4;	Length	551;
Best Local Similarity	88.0%;	Pred. No.	4,7e-236;				
Matches	485;	Conservative	35;	Mismatches	31;	Indels	0;
Gaps							
Qy	417	CASLNGVNDSHMASMLNLDHSPWSPSCSGYMITSLFDNGHGECMLDKPQNPQLPQD	476				
Db	1	CASLNGVSDSHLMASMLSSLDHSPWSPCSAYVYVTSFLFDNGHGECMLDKPQNPQLPQD	60				
Qy	477	LPGTSYDANRQCOPTFGEDSKHCPCDAASTCTSLTWCCTSGGVLVCQTKHFPWADGTSCE	536				
Db	61	LPGLTYDANRQCOPTFGESKHCPCDAASTCTTLCWCTSGGLVVCQTKHFPWADGTSCE	120				
Qy	537	GKWCINGKCVNKNHKKHFDTPPHGSGWMGMPGDCSRTCCTGGGVQVYTMRECDNPVPKNGK	596				
Db	121	GKWCYSGKCVNKTDMKHFPATPVHSGMPGMPGDCSRTCCTGGGVQVYTMRECDNPVPKNGK	180				
Qy	597	YCEGKRVYRSCNLDEDCPDNNKTFREOCEAHNEFASKASFGSGGPAVEWIPKYAGVSPKD	656				
Db	181	YCEGKRVYRSCNIEDCPDNNKTFREOCEAHNEFASKASFGNEFTVEWTPKYAGVSPKD	240				
Qy	657	RCKLICQAKGIGYFFVLQPVKVDGTPCSPDSTSVCVQGCYKAGGDR1IDSKKKFDKCV	716				

```

Db      241  RCKLTCEAKGIGYFVLPQKVVDTGTPCSPDSTSCVQGCVKAGCDRIIDSXKKFDKGV 300
Qy      717  CGNGSGTCKKISGVSWSAKPGYHDIITPTGATNIEVKQRNQRSGRNNGSFLAIKAADGT 776
Db      301  CGNGSGTCKKWSGIVSTRPGYHDIIVTPAGATNIEVKHRNQRSGRNNGSFLAIRAADGT 360
Qy      777  YILNGDYTLSTLEODIMYKGVVLYRSGSSAALERIRSFSPKLEPTIQVLTVGNALRPKI 836
Db      361  YILNGNFTLSTLEODLTGKTVLYRSGSSAALERIRSFSPKLEPTIQVLMVGHALRPKI 420
Qy      837  KYTYFVKKKESFNAIPTFSAWVLEEWCESCSCBELGQWRRLVECRDINGQPASECAKEV 896
Db      421  KFTYFMKKKTESFNAIPTFSEWVLEEWCESCSTCGSGQWRVVCQCRDINGHPASECAKEV 480
Qy      897  KPASTRPCADHPCPQWOLGESSCSKTCGKGYYKKTSLKCLSHDGGVLSHSDCDPLKKPKH 956
Db      481  KPASTRPCADLPCPHWQVGDWSPCSKTCGKGYYKKTSLKCLSHDGGVLSHSDCDPLKKPKH 540
Qy      957  FIDFCTMAECS 967
Db      541  YIDFCTLTQCS 551

RESULT 8
US-10-009-332-1
; Sequence 1, Application US/10009332
; Patent No. 6716613
; GENERAL INFORMATION:
; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
; APPLICANT: Kazusa DNA Research Institute
; TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING AGGREGANASE ACTIVITY
; FILE REFERENCE: Q67541
; CURRENT APPLICATION NUMBER: US/10/009,332
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: JPA Hei 11-321740
; PRIOR FILING DATE: 1999-11-11
; PRIOR APPLICATION NUMBER: JPA 2000-144020
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-009-332-1

```

Query Match	46.7%	Score	2471.5	DB 4	Length	950			
Best Local Similarity	48.8%	Pred.	No. 2.1e-207						
Matches	484	Conservative	153	Mismatches	253	Indels	101	Gaps	24
Qy	36	LLLLAALLAVSDALGRPSEDESLVP-ELE-----RVP-----GHGTRRLRLHAF	81						
Db	1	MLLLGILTLAFAGTAGGFEPEREVVPRIQLDPDINGRRYYWRGPEDSGDQGLIFQITAF	60						
Qy	82	DQQLDLVDPDSSFLAGPFTLQNVGRKSGSDTLP-----ETDLAHCFVSGTVNGDPSSA	136						
Db	61	QEDFYHLTLPDAQFLAPAFSTHLG-----VPLQGLTGGSSDLRRRCFYSYGDVNAEPDSF	114						
Qy	137	AAFLSLCGVGFAYLLGEAYFIQLPAASERLATAPGCEKPPAPLQFHLRLRNROGDVGG	196						
Db	115	AAVSLCGGLRGAFGYRGAEYVISPLPNAS-----APAAQNSOGA-----HLLQ--RRGVPGG	165						
Qy	197	TCGVVDDEPRPTGKAETEDEGTEGEDGPQWS-----PQDPALQGVGP--TGTGS	247						
Db	166	PSG-----DPTSC-----GVNSG--WNPAILRALDPYKPRRAGGESRRRS	207						
Qy	248	IRKKRFVSSHRYVBETMLVADQSMAEFHGSGLGHYLLTLFSVAARLYKHPISIRNSVLVV	307						
Db	208	GRAKRFVSIPIRYVETLVVADSMVKFHGADLEHYLLTLTLLATAARLYRHPISILNPINVV	267						
Qy	308	KILVIHDEQKGPVTSNAALTLRNFNCWQKHNPSPSDRAEHYDTAILFTTRQDLCSQTC	367						



; NUMBER OF SEQ ID NOS: 21  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 837  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-122-126B-2

QY	36	LLLLAAALLAVSDALGRPSEDEELVVP	EL---ERVPGHGT---TRL--RLHAFDQQLDLVD	89
Db	37	LLLLLLASLLPSARLASPLPREEBIVF	PEKLSGVLPGSGAPARLLCLQAFGETLLLEL	96
QY	90	PPDSFLAPGFTLQNVGRK----	SGSDTLPETDLAHCFYSGTVNGDPSSAALSICG-	144
Db	97	EODSGVQVEGLTVQLGQAPPELLGAE	---PGT-----YLTGTINGDPESVASLHWDGA	148
QY	145	VRGAFYLLGEAYFIQPLPAASERLATA	APGKPPAPLQFHLRLRRNQGDVGTCGVVDDE	204
Db	149	LLGVLYRGAEHLQPLEGGTNSA-GG	PGGA-----HILRRK-----	184
QY	205	PRPTGKAETEDDEGTEGEDGEPQSP	QDPALQGVQGTGSGIRKRFVSHRYVETML	264
Db	185	-----SPASGQGPCNVKAP----	LGSPSPRPR-RAKRFASLSRREVETLV	224
QY	265	VADQSMABFHGSLKHYLLTLFSVAAR	LYKHPSIRNSVSLVVKLVVHDEQKGPVTSN	324
Db	225	VADDKMAAFHAGLAKRYLLTVMAAA	AKAFKPSIRNPVSLVTVRLVILGSGEGEPQVGS	284
QY	325	AALTLRNFQKQHNPPSDRDAEHYDTA	ILFTTRQDLGCSQTCDTLGMADVGTVCDPSPS	384
Db	285	AAQTLSFCAMQRLNTPEDSDPHDFTA	ILFTTRQDLGCVSTCDTLGMADVGTVCDPARS	344
QY	385	CSVIEDGLQAAFTTAHELGHVFNPHD	DAKQASLNG-VNQDSHMMASMLSNLDHSPW	443
Db	345	CAIVEDDGLQSAFTAHAHELGHVFN	MLHDNSKPCISLNGPLSTSRHVMAVMAHVDPBPW	404
QY	444	SPCSGYMTITSLDNGHGECLMDKPN	QIQLPGDLPGTSDYDANROCOFTFGEDSKHCPDAA	503
Db	405	SPCSARFITDFLDNGYGHCLLDKPE	APLHLPTVTFPGKDYDADROCOLTTFGPDSSRHC	464
QY	504	STCSTLWCTGTSGGVLVQCOTKHFP	WADGTSCEGKWCINGKCVNKNHKKHFDTPFHG	563
Db	465	PPCAALWCSGHLNGHAMCOTKHSP	WADGTGCPGPAQACMGGRCLHMDQLQDFNI	524
QY	564	MWGPWGDCSRTCGGVQVYTMRECDN	PNVPKNGGKYCEGKRVYRSCNLEDCPDNNKGTRE	623
Db	525	PWGPWGDCSRTCGGVQVFSRRDCTR	PPVPRNGGKYCEGRRTFRSCNTEDCPTGSALTRE	584
QY	624	EOCEAHNEFSKASFCSGPA-VEMI	PKYAGVSPKDRCKLIQAKGIGYFVLQPKVVDGTP	682
Db	585	EQCAAYNHRTDL-FKSPFGPMDVPR	YTGVAPODQCKLTCQARALGYVYVLEPRVVDGTP	643
QY	683	CSPDSTSVCOQCQVYKAGCDRIID	SKKFKDCKGCGGNGSTCKKISGVS TSAKPGYHDI	742
Db	644	CSPDSSSVCOQRCIHAGCDRIIG	SKKFKDCKWCGGSGCSKQSGSFRKFRYGYNNV	703
QY	743	TIPTGATNIEVKQRNQRNNGSFLA	IKAADGTYILNGDYTLSTLEODIMYKGVV-LRY	801
Db	704	TIPAGATHILVRQGNPGHRS--IY	LALKLPDGSYALNGEYTLMPSPDTPVLP	761
QY	802	SGSSAALIRIRFSPLEPLTIQVLT	VGNALRPKIKYTYFVKKKKESNAIPTSAW	859
Db	762	SGATAASETLGSHGFLAQPLTLQ	VLVAGNPQDTRLRYSFVFRPTPS--TPRPTQD	818

RESULT 11

US-09-634-286A-2  
 ; Sequence 2, Application US/09634286A  
 ; Patent No. 6521436

; GENERAL INFORMATION:  
 ; APPLICANT: Bristol-Myers Squibb Company  
 ; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES  
 ; FILE REFERENCE: DM6909A  
 ; CURRENT APPLICATION NUMBER: US/09/634,286A  
 ; CURRENT FILING DATE: 2000-08-09  
 ; NUMBER OF SEQ ID NOS: 21  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 837  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-634-286A-2

Query Match	40.2%;	Score 2124;	DB 4;	Length 837;
Best Local Similarity	49.0%;	Pred. No. 5e-177;		
Matches 411;	Conservative 132;	Mismatches 225;	Indels 70;	Gaps 18;
QY	36	LLLLAAALLAVSDALGRPSEDEELVVP	EL---ERVPGHGT---TRL--RLHAFDQQLDLVD	89
Db	37	LLLLLLASLLPSARLASPLPREEBIVF	PEKLSGVLPGSGAPARLLCLQAFGETLLLEL	96
QY	90	PPDSFLAPGFTLQNVGRK----	SGSDTLPETDLAHCFYSGTVNGDPSSAALSICG-	144
Db	97	EODSGVQVEGLTVQLGQAPPELLGAE	---PGT-----YLTGTINGDPESVASLHWDGA	148
QY	145	VRGAFYLLGEAYFIQPLPAASERLATA	APGKPPAPLQFHLRLRRNQGDVGTCGVVDDE	204
Db	149	LLGVLYRGAEHLQPLEGGTNSA-GG	PGGA-----HILRRK-----	184
QY	205	PRPTGKAETEDDEGTEGEDGEPQSP	QDPALQGVQGTGSGIRKRFVSHRYVETML	264
Db	185	-----SPASGQGPCNVKAP----	LGSPSPRPR-RAKRFASLSRREVETLV	224
QY	265	VADQSMABFHGSLKHYLLTLFSVAAR	LYKHPSIRNSVSLVVKLVVHDEQKGPVTSN	324
Db	225	VADDKMAAFHAGLAKRYLLTVMAAA	AKAFKPSIRNPVSLVTVRLVILGSGEGEPQVGS	284
QY	325	AALTLRNFQKQHNPPSDRDAEHYDTA	ILFTTRQDLGCSQTCDTLGMADVGTVCDPSPS	384
Db	285	AAQTLSFCAMQRLNTPEDSDPHDFTA	ILFTTRQDLGCVSTCDTLGMADVGTVCDPARS	344
QY	385	CSVIEDGLQAAFTTAHELGHVFNPHD	DAKQASLNG-VNQDSHMMASMLSNLDHSPW	443
Db	345	CAIVEDDGLQSAFTAHAHELGHVFN	MLHDNSKPCISLNGPLSTSRHVMAVMAHVDPBPW	404
QY	444	SPCSGYMTITSLDNGHGECLMDKPN	QIQLPGDLPGTSDYDANROCOFTFGEDSKHCPDAA	503
Db	405	SPCSARFITDFLDNGYGHCLLDKPE	APLHLPTVTFPGKDYDADROCOLTTFGPDSSRHC	464
QY	504	STCSTLWCTGTSGGVLVQCOTKHFP	WADGTSCEGKWCINGKCVNKNHKKHFDTPFHG	563
Db	465	PPCAALWCSGHLNGHAMCOTKHSP	WADGTGCPGPAQACMGGRCLHMDQLQDFNI	524
QY	564	MWGPWGDCSRTCGGVQVYTMRECDN	PNVPKNGGKYCEGKRVYRSCNLEDCPDNNKGTRE	623
Db	525	PWGPWGDCSRTCGGVQVFSRRDCTR	PPVPRNGGKYCEGRRTFRSCNTEDCPTGSALTRE	584
QY	624	EOCEAHNEFSKASFCSGPA-VEMI	PKYAGVSPKDRCKLIQAKGIGYFVLQPKVVDGTP	682
Db	585	EQCAAYNHRTDL-FKSPFGPMDVPR	YTGVAPODQCKLTCQARALGYVYVLEPRVVDGTP	643
QY	683	CSPDSTSVCOQCQVYKAGCDRIID	SKKFKDCKGCGGNGSTCKKISGVS TSAKPGYHDI	742
Db	644	CSPDSSSVCOQRCIHAGCDRIIG	SKKFKDCKWCGGSGCSKQSGSFRKFRYGYNNV	703
QY	743	TIPTGATNIEVKQRNQRNNGSFLA	IKAADGTYILNGDYTLSTLEODIMYKGVV-LRY	801
Db	704	TIPAGATHILVRQGNPGHRS--IY	LALKLPDGSYALNGEYTLMPSPDTPVLP	761
QY	802	SGSSAALIRIRFSPLEPLTIQVLT	VGNALRPKIKYTYFVKKKKESNAIPTSAW	859
Db	762	SGATAASETLGSHGFLAQPLTLQ	VLVAGNPQDTRLRYSFVFRPTPS--TPRPTQD	818





QY 624 EQCAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLIQAKGIGYFFVLQPKVVDGTPC 683  
 DB 631 EQCEAKNGYQSDAKGVKTFVWPKYAGVLPADVCKLTCRAKGTGYVVFSPKVTGTEC 690  
 QY 684 SPDSTSVCOGCVKAGCDRIIDSKKPKKCGVCGNGSTCKKISGVSATKPGYHDIIT 743  
 DB 691 RPYNSVVRGKCVRTGCDGIIGSKLQYDKGCVCGDNSSCTKIIVTFNKKSKGYTDVVR 750  
 QY 744 IPTGATNIEVKORNGSRNNGSFLAIKAADGTIYILNGDYTLSTLEODIMYKGVVLYRSG 803  
 DB 751 IPEGATHIKVRQPKAKQDRTFTAYLALKKNGEYLINGKYMISTETIIDINGTVMNYS 810  
 QY 804 SSAALERIR--SFSPKLEPLTIQVLTVGNALRPKIYTFV--KKKESFNAIPTFSA--- 857  
 DB 811 WSHRDDFLHGMYSATKEILIVQILATDPTKLDVRYSPFVKSPKNSVTSHGSKNV 870  
 QY 858 -----WVIERWGECSKCELGQWRRLVECRDINGQPASECAKEVPASTRPCADHPC 909  
 DB 871 GSHTSQPWVTGPWLACSTRCTDGTGWHTRTVQCDGNRKLAKGCPLSQRPSPAFKQCLLKKC 930  
 RESULT 14  
 US-09-634-286A-15  
 ; Sequence 15, Application US/09634286A  
 ; Patent No. 6521436  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bristol-Myers Squibb Company  
 ; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES  
 ; FILE REFERENCE: DM6909A  
 ; CURRENT APPLICATION NUMBER: US/09/634, 286A  
 ; CURRENT FILING DATE: 2000-08-09  
 ; NUMBER OF SEQ ID NOS: 21  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 15  
 ; LENGTH: 930  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-634-286A-15  
 \* Query Match 36.4%; Score 1924; DB 4; Length 930;  
 Best Local Similarity 40.7%; Pred. No. 2e-159;  
 Matches 391; Conservative 125; Mismatches 314; Indels 130; Gaps 17;  
 QY 25 PGSRFGPVPPTLLAAALLAVSDALGRPSEDEELVPELERVPGHG---TTLRLRHAF 81  
 DB 26 PAQDKAGQPPT-----AAAAQPRRQGEVEQERAEP-----PGHPLAQRRRSKGL 73  
 QY 82 DQQLD-----LDVPPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120  
 DB 74 VQNIQOLYSGGKGVLYVAGRRFLDLDERDGSVGIAGFV-----PAGGTSAPMRHR 127  
 QY 121 AHCFYSGTVNGDPPSSAALSICEVGRGAFYLLGEAYFIQPL---PAASER----- 167  
 DB 128 SHCFYRGVTVDASPRSLAVFDLCGGLDGFFAVKHARYTLKPLLRGPAWEEKGRVYDGS 187  
 QY 168 -----LATAAPEKPPAPLQFHLRRNRQDVGCTGVVDDEP---RPTG 209  
 DB 188 RILHYVTRREGFSFEALPPRASCETPASTPEAH-----EHAPAHNSP 230  
 QY 210 KA-----ETEBDEGTEGEDEGPQ--WSPQDPALQGVQPTGTGSIIRKRFVSSHRYVETM 263  
 DB 231 RAALASQLLDQALSAPGSGPQTWNR-----RRRSISRARQVEILL 272  
 QY 264 LVADQSMAPFHSGSLKHYLLTLFSVAARLYKHSIRNSVLSVAVVLLVHDEQKGPVTS 323  
 DB 273 LVADASMARLYGRGLQHYLLTLASTANRLYSHASINHLRAWVWVVLGDKDKSLEVS 332  
 QY 324 NAALTILRNFQKQHNPPSDRAEDHYDTAILFTRODLGSGTCTDLTGADVGTVCDPDR 383  
 DB 333 NAATTLKFNKQKQHNQNGDDHEHYDAIILFTREDLCGHHSCTDLTGADVGTTCSPER 392  
 QY 384 SCSVIEDDGLQAAFTTAHBLGHVFNMPHDDAKQOCASLNGVNDSHMMSMLNLDHSPW 443

DB 393 SCVIEDDGLHAAFTVAHEIIGHLGLSHDSDSKFCEETFGSTEDKRLMSSILTSIDASKPW 452  
 QY 444 SPSCGYMITTSLDNGHGBCLMDKPNPTQLPGDLPPTSVDANROCOFTFGSDSKHCPDAA 503  
 DB 453 SKCTSATITFELDDHGNCLLDLPKQILGPEELPGQTYDATQCNLTFFGEYSVCP--GM 511  
 QY 504 STCSTLWCTGTSGGVLVCQTKHFPWADGTSCEGKWCINGKCVNKNHRRKHFDTPPHGWS 563  
 DB 512 DVCARLWCAVVRQGMVCLTKKLPAVEGTPCGKGRICLQKGCVDKTKKKYISTSSHGNG 571  
 QY 564 MWGPWGDSCRPCGGVQYTMRECDNPVPAKGGKYCEGKRVRYRSCNLEDCPDNNGKTPRE 623  
 DB 572 SWGSGWQCSRCGGGVQPAYRHCNNPAPRNNRGVCTGKRAIYRSCSLMPCPP--NGKSPRH 630  
 QY 624 EQCAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLIQAKGIGYFFVLQPKVVDGTPC 683  
 DB 631 EQCEAKNGYQSDAKGVKTFVWPKYAGVLPADVCKLTCRAKGTGYVVFSPKVTGTEC 690  
 QY 684 SPDSTSVCOGCVKAGCDRIIDSKKPKKCGVCGNGSTCKKISGVSATKPGYHDIIT 743  
 DB 691 RPYNSVVRGKCVRTGCDGIIGSKLQYDKGCVCGDNSSCTKIIVTFNKKSKGYTDVVR 750  
 QY 744 IPTGATNIEVKORNGSRNNGSFLAIKAADGTIYILNGDYTLSTLEODIMYKGVVLYRSG 803  
 DB 751 IPEGATHIKVRQPKAKQDRTFTAYLALKKNGEYLINGKYMISTETIIDINGTVMNYS 810  
 QY 804 SSAALERIR--SFSPKLEPLTIQVLTVGNALRPKIYTFV--KKKESFNAIPTFSA--- 857  
 DB 811 WSHRDDFLHGMYSATKEILIVQILATDPTKLDVRYSPFVKSPKNSVTSHGSKNV 870  
 QY 858 -----WVIERWGECSKCELGQWRRLVECRDINGQPASECAKEVPASTRPCADHPC 909  
 DB 871 GSHTSQPWVTGPWLACSTRCTDGTGWHTRTVQCDGNRKLAKGCPLSQRPSPAFKQCLLKKC 930  
 RESULT 15  
 US-10-247-685-15  
 ; Sequence 15, Application US/10247685  
 ; Patent No. 6753176  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bristol-Myers Squibb Company  
 ; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES  
 ; FILE REFERENCE: DM6909D  
 ; CURRENT APPLICATION NUMBER: US/10/247,685  
 ; CURRENT FILING DATE: 2002-09-19  
 ; NUMBER OF SEQ ID NOS: 21  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 15  
 ; LENGTH: 930  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-247-685-15  
 Query Match 36.4%; Score 1924; DB 4; Length 930;  
 Best Local Similarity 40.7%; Pred. No. 2e-159;  
 Matches 391; Conservative 125; Mismatches 314; Indels 130; Gaps 17;  
 QY 25 PGSRFGPVPPTLLAAALLAVSDALGRPSEDEELVPELERVPGHG---TTLRLRHAF 81  
 DB 26 PAQDKAGQPPT-----AAAAQPRRQGEVEQERAEP-----PGHPLAQRRRSKGL 73  
 QY 82 DQQLD-----LDVPPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120  
 DB 74 VQNIQOLYSGGKGVLYVAGRRFLDLDERDGSVGIAGFV-----PAGGTSAPMRHR 127  
 QY 121 AHCFYSGTVNGDPPSSAALSICEVGRGAFYLLGEAYFIQPL---PAASER----- 167  
 DB 128 SHCFYRGVTVDASPRSLAVFDLCGGLDGFFAVKHARYTLKPLLRGPAWEEKGRVYDGS 187  
 QY 168 -----LATAAPEKPPAPLQFHLRRNRQDVGCTGVVDDEP---RPTG 209  
 DB 188 RILHYVTRREGFSFEALPPRASCETPASTPEAH-----EHAPAHNSP 230



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2005, 10:12:49 ; Search time 167 Seconds  
(without alignments)  
2267.448 Million cell updates/sec

Title: US-09-989-687-126  
Perfect score: 5287  
Sequence: 1 MORAVPEGFGRKLGSDMGN.....CDPLKKPKHFIDFCTMASC 967

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 1759131

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5287	100.0	967	11	US-09-989-687-126
2	5232	99.0	967	13	US-10-105-929-2
3	5232	99.0	967	14	US-10-115-286-2
4	5232	99.0	967	16	US-10-757-450-2
5	5234	98.8	967	16	US-10-755-889-134
6	5224	98.8	967	17	US-10-741-600-1603
7	5224	98.8	967	17	US-10-741-600-1604
8	5224	98.8	967	18	US-10-923-035-37
9	5221.5	98.8	968	10	US-09-373-658-125
10	5144	97.3	950	10	US-09-373-658-2
11	5144	97.3	950	11	US-09-989-687-2

12	5140	97.2	949	17	US-10-667-281-2	Sequence 2, Appli
13	5136	97.1	950	17	US-10-741-600-1605	Sequence 1605, Ap
14	5136	97.1	950	18	US-10-973-858-14	Sequence 14, Appl
15	5036	95.3	931	9	US-09-741-151-4	Sequence 4, Appli
16	4293.5	81.2	968	13	US-10-163-316-7	Sequence 7, Appli
17	4293.5	81.2	968	15	US-10-391-364-82	Sequence 82, Appl
18	4264.5	80.7	951	15	US-10-381-793-3	Sequence 3, Appli
19	4260.5	80.6	950	9	US-09-321-987B-4	Sequence 4, Appli
20	3922	74.2	727	9	US-09-445-023A-1	Sequence 1, Appli
21	3922	74.2	727	14	US-10-037-597-1	Sequence 1, Appli
22	3922	74.2	727	14	US-10-037-580-1	Sequence 12, Appl
23	3656	69.2	727	9	US-09-445-023A-12	Sequence 12, Appl
24	3656	69.2	727	14	US-10-037-597-12	Sequence 12, Appl
25	3656	69.2	727	14	US-10-037-580-12	Sequence 2, Appli
26	3303	62.5	608	9	US-09-803-589-2	Sequence 8, Appli
27	3303	62.5	608	16	US-10-718-332-2	Sequence 2, Appli
28	3297	62.4	608	13	US-09-803-589-8	Sequence 13, Appl
29	3297	62.4	608	13	US-10-105-929-13	Sequence 8, Appli
30	3297	62.4	608	16	US-10-718-332-8	Sequence 16, Appl
31	2794	52.8	551	9	US-09-802-582-16	Sequence 16, Appl
32	2794	52.8	551	13	US-10-105-929-16	Sequence 16, Appl
33	2794	52.8	551	14	US-10-365-227-16	Sequence 10, Appl
34	2635	49.8	518	9	US-09-803-589-10	Sequence 10, Appl
35	2635	49.8	518	16	US-10-718-332-10	Sequence 10, Appl
36	2495.5	47.2	924	15	US-10-032-463-28	Sequence 28, Appl
37	2478	46.9	978	15	US-10-275-107-59	Sequence 59, Appl
38	2477.5	46.9	950	9	US-09-741-151-2	Sequence 2, Appli
39	2477.5	46.9	950	16	US-09-965-631-4	Sequence 4, Appli
40	2477.5	46.9	950	17	US-10-391-364-77	Sequence 77, Appl
41	2477.5	46.9	950	17	US-10-753-267-56	Sequence 56, Appl
42	2477.5	46.9	950	17	US-10-961-020-4	Sequence 4, Appli
43	2471.5	46.7	950	16	US-10-763-210-1	Sequence 1, Appli
44	2467	46.7	577	15	US-10-425-114-39241	Sequence 39241, A
45	2455.5	46.4	952	15	US-10-311-035-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1  
US-09-989-687-126  
; Sequence 126, Application US/0989687  
; Publication No. US20040002449A1  
; GENERAL INFORMATION:  
; APPLICANT: Hastings, Gregg A.  
; APPLICANT: Ruben, Steven M.  
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides  
; FILE REFERENCE: 1488.107000D  
; CURRENT APPLICATION NUMBER: US/09/989,687  
; CURRENT FILING DATE: 2001-11-21  
; NUMBER OF SEQ ID NOS: 126  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 126  
; LENGTH: 967  
; TYPE: PRT  
; ORGANISM: ITGL-TSP  
US-09-989-687-126

Query Match 100.0%; Score 5287; DB 11; Length 967;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 967; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MORAVPEGFGRKLGSDMGNAPGSRSGFVPPTLLLAALLAVSDALGRPSEDEEL	60
Db	1	MORAVPEGFGRKLGSDMGNAPGSRSGFVPPTLLLAALLAVSDALGRPSEDEEL	60
Qy	61	VPELVRVPGHGTTRLRHAFDQQLDLPDPSFLAGFTLQNVGRKSGSDTLPETDL	120
Db	61	VPELVRVPGHGTTRLRHAFDQQLDLPDPSFLAGFTLQNVGRKSGSDTLPETDL	120
Qy	121	AHCFYSGTVNGDPPSSAAALSCEGVGFAYLLGEAYFIQPLPAASERLATAPGKPPAP	180
Db	121	AHCFYSGTVNGDPPSSAAALSCEGVGFAYLLGEAYFIQPLPAASERLATAPGKPPAP	180

```
QY 181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKARTEDDEDETEDEGPGQSPDPAALQGVG 240
Db 181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKARTEDDEDETEDEGPGQSPDPAALQGVG 240
QY 241 QPTGTGSIKRRKRFVSSHRYVETMLVADQSMAEFFHSGSLKHYLLTLFVSAARLYKHPISRN 300
Db 241 QPTGTGSIKRRKRFVSSHRYVETMLVADQSMAEFFHSGSLKHYLLTLFVSAARLYKHPISRN 300
QY 301 SVSLVVVILVHIDRQKPEVTSNAALTLRPNFCWQKQHNPPSDRDAEHYDTAILFTTQD 360
Db 301 SVSLVVVILVHIDRQKPEVTSNAALTLRPNFCWQKQHNPPSDRDAEHYDTAILFTTQD 360
QY 361 LCGSQTCDTLGMADVGTVCDSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
Db 361 LCGSQTCDTLGMADVGTVCDSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
QY 421 NGVNQDSHMMASMLNLDHSPWSPCSGYMITSLFDNGHGCECLMDKPNPQLPGDLPGT 480
Db 421 NGVNQDSHMMASMLNLDHSPWSPCSGYMITSLFDNGHGCECLMDKPNPQLPGDLPGT 480
QY 481 SYDANRQCQFTFGEDSKHCIPDAASTCSTLWCTGTSGGVLVCQTKHFPWADGTSCEGKWC 540
Db 481 SYDANRQCQFTFGEDSKHCIPDAASTCSTLWCTGTSGGVLVCQTKHFPWADGTSCEGKWC 540
QY 541 INKCVNKNHRKHFDTPFHGSGWGMWGPWGDCSRTCGGGVQYTMRECDNPVPKNGKCYCEG 600
Db 541 INKCVNKNHRKHFDTPFHGSGWGMWGPWGDCSRTCGGGVQYTMRECDNPVPKNGKCYCEG 600
QY 601 KRVRYSNLEDCPDNNGKTFREEQCEAHNEFASKASFGSGPAVEWIPKYAGVSPKDRCKL 660
Db 601 KRVRYSNLEDCPDNNGKTFREEQCEAHNEFASKASFGSGPAVEWIPKYAGVSPKDRCKL 660
QY 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVQVQGVKAGCDRIIDSKKKFKDKGVCVCGN 720
Db 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVQVQGVKAGCDRIIDSKKKFKDKGVCVCGN 720
QY 721 GSTCKKISGSVTSKAPGYHDIITPTGATNIEVKORNGSRNNGSFIAIKAADGTYLILN 780
Db 721 GSTCKKISGSVTSKAPGYHDIITPTGATNIEVKORNGSRNNGSFIAIKAADGTYLILN 780
QY 781 GDYTLSTLEQDIIMYKGVVLYRSGSAALERIRSFPLKEPLTIQVLTGVALRPKIKYTY 840
Db 781 GDYTLSTLEQDIIMYKGVVLYRSGSAALERIRSFPLKEPLTIQVLTGVALRPKIKYTY 840
QY 841 FVKKKESFNAIPTFSAWVIEEWGECSSKSELGWQRRLVECRDINGQPASECAKEVKPAS 900
Db 841 FVKKKESFNAIPTFSAWVIEEWGECSSKSELGWQRRLVECRDINGQPASECAKEVKPAS 900
QY 901 TRPCADHPCPQWOLGEWSSCKTCGKYKTKSLKCLSHDGGVLSDSCDPLKKPKHFDIF 960
Db 901 TRPCADHPCPQWOLGEWSSCKTCGKYKTKSLKCLSHDGGVLSDSCDPLKKPKHFDIF 960
QY 961 CTMAECS 967
Db 961 CTMAECS 967
```

## RESULT 2

```
US-10-105-929-2
; Sequence 2, Application US/10105929
; Publication No US20020137142A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/10/105,929
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/130,491
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-05
```

```
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/054,961
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-105-929-2
```

```
Query Match 99.0%; Score 5232; DB 13; Length 967;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 957; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 MORAVPFGFGRKLGSDMGNAERAPGRSPGVPVTLALLAAALAVDALGRPSEDEEL 60
Db 1 MORAVPFGFGRKLGSDMGNAERAPGRSPGVPVTLALLAAALAVDALGRPSEDEEL 60
QY 61 VVPELERVPVGHGTTRLRLHAFDQQLDLVPPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
Db 61 VVPELERVPVGHGTTRLRLHAFDQQLDLVPPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
QY 121 AHCFSYGTNVGDPSSAAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKEPPAP 180
Db 121 AHCFSYGTNVGDPSSAAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKEPPAP 180
QY 181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKARTEDDEDETEDEGPGQSPDPAALQGVG 240
Db 181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKARTEDDEDETEDEGPGQSPDPAALQGVG 240
QY 241 QPTGTGSIKRRKRFVSSHRYVETMLVADQSMAEFFHSGSLKHYLLTLFVSAARLYKHPISRN 300
Db 241 QPTGTGSIKRRKRFVSSHRYVETMLVADQSMAEFFHSGSLKHYLLTLFVSAARLYKHPISRN 300
QY 301 SVSLVVVILVHIDRQKPEVTSNAALTLRPNFCWQKQHNPPSDRDAEHYDTAILFTTQD 360
Db 301 SVSLVVVILVHIDRQKPEVTSNAALTLRPNFCWQKQHNPPSDRDAEHYDTAILFTTQD 360
QY 361 LCGSQTCDTLGMADVGTVCDSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
Db 361 LCGSQTCDTLGMADVGTVCDSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
QY 421 NGVNQDSHMMASMLNLDHSPWSPCSGYMITSLFDNGHGCECLMDKPNPQLPGDLPGT 480
Db 421 NGVNQDSHMMASMLNLDHSPWSPCSGYMITSLFDNGHGCECLMDKPNPQLPGDLPGT 480
QY 481 SYDANRQCQFTFGEDSKHCIPDAASTCSTLWCTGTSGGVLVCQTKHFPWADGTSCEGKWC 540
Db 481 SYDANRQCQFTFGEDSKHCIPDAASTCSTLWCTGTSGGVLVCQTKHFPWADGTSCEGKWC 540
QY 541 INKCVNKNHRKHFDTPFHGSGWGMWGPWGDCSRTCGGGVQYTMRECDNPVPKNGKCYCEG 600
Db 541 INKCVNKNHRKHFDTPFHGSGWGMWGPWGDCSRTCGGGVQYTMRECDNPVPKNGKCYCEG 600
QY 601 KRVRYSNLEDCPDNNGKTFREEQCEAHNEFASKASFGSGPAVEWIPKYAGVSPKDRCKL 660
Db 601 KRVRYSNLEDCPDNNGKTFREEQCEAHNEFASKASFGSGPAVEWIPKYAGVSPKDRCKL 660
QY 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVQVQGVKAGCDRIIDSKKKFKDKGVCVCGN 720
Db 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVQVQGVKAGCDRIIDSKKKFKDKGVCVCGN 720
QY 721 GSTCKKISGSVTSKAPGYHDIITPTGATNIEVKORNGSRNNGSFIAIKAADGTYLILN 780
Db 721 GSTCKKISGSVTSKAPGYHDIITPTGATNIEVKORNGSRNNGSFIAIKAADGTYLILN 780
QY 781 GDYTLSTLEQDIIMYKGVVLYRSGSAALERIRSFPLKEPLTIQVLTGVALRPKIKYTY 840
Db 781 GDYTLSTLEQDIIMYKGVVLYRSGSAALERIRSFPLKEPLTIQVLTGVALRPKIKYTY 840
QY 841 FVKKKESFNAIPTFSAWVIEEWGECSSKSELGWQRRLVECRDINGQPASECAKEVKPAS 900
Db 841 FVKKKESFNAIPTFSAWVIEEWGECSSKSELGWQRRLVECRDINGQPASECAKEVKPAS 900
```

QY 901 TRPCADHPCPQWLGWSSCKTCGKYKTSKLCSLSHDGGVLSHSDCDPLKKPKHFIDF 960  
Db 901 TRPCADHPCPQWLGWSSCKTCGKYKTSKLCSLSHDGGVLSHSDCDPLKKPKHFIDF 960  
QY 961 CTMAECS 967  
Db 961 CTMAECS 967

RESULT 3  
US-10-115-286-2  
; Sequence 2, Application US/10115286  
; Publication No. US20030166065A1  
; GENERAL INFORMATION:  
; APPLICANT: Jonak, Zdenka  
; Trulli, Stephen  
; Fornwald, James  
; Terrett, Jonathan  
; Hastings, Gregg  
; TITLE OF INVENTION: No. US20030166065A1e1 Integrin Ligand ITGL-TSP  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ratner & Prestia  
; STREET: Box 980  
; CITY: Valley Forge  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19482

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/115,286  
; FILING DATE: 04-Apr-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/845,496  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Prestia, Paul F.  
; REGISTRATION NUMBER: 23,031  
; REFERENCE/DOCKET NUMBER: GH-70000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-407-0700  
; TELEFAX: 610-407-0701  
; TELEX: 846169

INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 967 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-115-286-2

Query Match 99.0%; Score 5232; DB 14; Length 967;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 957; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 MORAVPEGGRKLGSDMGNAERAPGSRFGVPVPTLLLLAALLAVSDALGRPSEDEEL 60  
Db 1 MORAVPEGGRKLGSDMGNAERAPGSRFGVPVPTLLLLAALLAVSDALGRPSEDEEL 60  
QY 61 VVPELVRPGHGTTRLRHAFQQLDLOVPPDSFSLAPGFTLLQNVGRKSGSDTLPETDL 120  
Db 61 VVPELERAPGHGTTRLRHAFQQLDLELRPDSSFLAPGFTLLQNVGRKSGSETPLPETDL 120  
QY 121 AHCFYSGTVNGDPSAAALSLCEGVRGAFYLLGBAYFTIQLPLPAASERLATAPGKPPAP 180

Db 121 AHCFYSGTVNGDPSAAALSLCEGVRGAFYLLGBAYFTIQLPLPAASERLATAPGKPPAP 180  
QY 181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETDEDEGEDEGGPQWSPQDPALQGVG 240  
Db 181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETDEDEGEDEGGPQWSPQDPALQGVG 240  
QY 241 OPTGTGSIIRKKRFVSSSHRYVETMLVADQSMAEFHGSLGKHYLLTLFSAARLYKHPSIRN 300  
Db 241 OPTGTGSIIRKKRFVSSSHRYVETMLVADQSMAEFHGSLGKHYLLTLFSAARLYKHPSIRN 300  
QY 301 SVSLVVVKILVIHDEOKGPEVTSNAALTLRNFCWKQKOHNPSPDRDAEHYDTAILFTTQD 360  
Db 301 SVSLVVVKILVIHDEOKGPEVTSNAALTLRNFCWKQKOHNPSPDRDAEHYDTAILFTTQD 360  
QY 361 LCGSQTCDTLGMADVGTVCDFPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACASL 420  
Db 361 LCGSQTCDTLGMADVGTVCDFPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACASL 420  
QY 421 NGVNQDSHMMASMLSLNLDHSPQWSPSCGYMTTSFLDNHGHGCLMDKQNPITQLPGDLPGT 480  
Db 421 NGVNQDSHMMASMLSLNLDHSPQWSPSCGYMTTSFLDNHGHGCLMDKQNPITQLPGDLPGT 480  
QY 481 SYDANRQCQFTFGEDSKHCPDAASTCSTLWCTGTSGGVLVCQTKHFPWADGTSCGEGKWC 540  
Db 481 SYDANRQCQFTFGEDSKHCPDAASTCSTLWCTGTSGGVLVCQTKHFPWADGTSCGEGKWC 540  
QY 541 INKCVNKNHRKHFDTPPFHGSMGMWPGDCSRTCCGGGVQYTMRECDNPPVKNKGKCYEG 600  
Db 541 INKCVNKNTRKHFDTPPFHGSMGMWPGDCSRTCCGGGVQYTMRECDNPPVKNKGKCYEG 600  
QY 601 KRVYRSCNLEDCPDNNGKTPREOCEAHNPFSAFSGSGPAVEWI PKYAGVSPKDRCKL 660  
Db 601 KRVYRSCNLEDCPDNNGKTPREOCEAHNPFSAFSGSGPAVEWI PKYAGVSPKDRCKL 660  
QY 661 ICOAKGIGYFFVLQPKVVDGTPCSPDSTSCVQCGVKAGCDRIIDSKKKFKDCKGCGGN 720  
Db 661 ICOAKGIGYFFVLQPKVVDGTPCSPDSTSCVQCGVKAGCDRIIDSKKKFKDCKGCGGN 720  
QY 721 GSTCKKISGSVTSAPGYVDIITPTGATNIEVQRNQRSGRNNGSFLAIKAADGTYILN 780  
Db 721 GSTCKKISGSVTSAPGYVDIITPTGATNIEVQRNQRSGRNNGSFLAIKAADGTYILN 780  
QY 781 GDYTLSTLEQIMYKGVVLYSGSSAALERIRSPKLEPTIOVLTVGNALRPKIKYTY 840  
Db 781 GDYTLSTLEQIMYKGVVLYSGSSAALERIRSPKLEPTIOVLTVGNALRPKIKYTY 840  
QY 841 FVKKKKESFNAIPTFSAWVIEEWGECSCSCLGWRRLVECRDINGOPASCAKEVKPAS 900  
Db 841 FVKKKKESFNAIPTFSAWVIEEWGECSCSCLGWRRLVECRDINGOPASCAKEVKPAS 900  
QY 901 TRPCADHPCPQWLGWSSCKTCGKYKTSKLCSLSHDGGVLSHSDCDPLKKPKHFIDF 960  
Db 901 TRPCADHPCPQWLGWSSCKTCGKYKTSKLCSLSHDGGVLSHSDCDPLKKPKHFIDF 960  
QY 961 CTMAECS 967  
Db 961 CTMAECS 967

RESULT 4  
US-10-757-450-2  
; Sequence 2, Application US/10757450  
; Publication No. US20040175794A1  
; GENERAL INFORMATION:  
; APPLICANT: Jonak, Zdenka  
; Trulli, Stephen  
; Fornwald, James  
; Terrett, Jonathan  
; Hastings, Gregg  
; TITLE OF INVENTION: Novel Integrin Ligand ITGL-TSP  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ratner & Prestia

```
STREET: Box 980
CITY: Valley Forge
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/10/757,450
  FILING DATE: 15-Jan-2004
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/10/115,286
  FILING DATE: 04-Apr-2002
  APPLICATION NUMBER: 08/845,496
  FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
  NAME: Prestia, Paul F.
  REGISTRATION NUMBER: 23,031
  REFERENCE/DOCKET NUMBER: GH-70000
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 610-407-0700
    TELEFAX: 610-407-0701
    TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
  LENGTH: 967 amino acids
  TYPE: amino acid
  STRANDEDNESS: single
  TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-757-450-2

Query Match          99.0%; Score 5232; DB 16; Length 967;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 957; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MQRVPEGFGRKRLGSDMGNAERAPGSRSGFVPPTLLLLAAALLAVSDALGRPSEDEEL 60
Db 1 MQRVPEGFGRKRLGSDMGNAERAPGSRSGFVPPTLLLLAAALLAVSDALGRPSEDEEL 60
Qy 61 VPELERVPGHGTTRLRHLHAFDQQLDLDVPPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
Db 61 VPELERAPGHGTTRLRHLHAFDQQLDLELRDPSFLAPGFTLQNVGRKSGSETLPETDL 120
Qy 121 AHCFYSGTVNGDPSSAAALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAP 180
Db 121 AHCFYSGTVNGDPSSAAALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAP 180
Qy 181 LQFHLLRRNRQDVGTCGVVDDEPRPTGKAETDEDEGEDEGEDEGEDEGEDEGEDEGEDE 240
Db 181 LQFHLLRRNRQDVGTCGVVDDEPRPTGKAETDEDEGEDEGEDEGEDEGEDEGEDEGEDE 240
Qy 241 OPTGTGSRKKRFVSSHRYVETMLVADQSMAEFGHSGLKHVLLTLFVVAARLYKHPISRN 300
Db 241 OPTGTGSRKKRFVSSHRYVETMLVADQSMAEFGHSGLKHVLLTLFVVAARLYKHPISRN 300
Qy 301 SVSLVVKVILVHDEQKGPVTSNAALTLRNFCNWKQHNPPSDRDAEHYDTAILFTROD 360
Db 301 SVSLVVKVILVHDEQKGPVTSNAALTLRNFCNWKQHNPPSDRDAEHYDTAILFTROD 360
Qy 361 LCGSQTCDTLGMADVTCVDPSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
Db 361 LCGSQTCDTLGMADVTCVDPSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
Qy 421 NGVNQDSHMMASMLNSLDHSPQSCGYMLTSLDNGHGECLMDKQNPLOLPGLDLPGT 480
Db 421 NGVNQDSHMMASMLNSLDHSPQSCGYMLTSLDNGHGECLMDKQNPLOLPGLDLPGT 480

103 SYDANROCOFTFGEDSKHCPDAASTCTLWCTGTSGGVLVCQTQKHPWADGTCGEGKWC 540
104 SYDANROCOFTFGEDSKHCPDAASTCTLWCTGTSGGVLVCQTQKHPWADGTCGEGKWC 540
105 INKCVNKNHRKHFDTPFHGSGWGMGPWDCSRTCCGGGVQVYTMRECDNPVPKNGGKYCEG 600
106 INKCVNKNTRKHFDTPFHGSGWGMGPWDCSRTCCGGGVQVYTMRECDNPVPKNGGKYCEG 600
107 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSAKSGSPGPAVEWIPKYAGVSPKDRCKL 660
108 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSAKSGSPGPAVEWIPKYAGVSPKDRCKL 660
109 ICQAKGIGYFVLQPKVVDGTPCSPDSTSVCOQCCKAGCDRIIDSKKFDKCGVCGGN 720
110 ICQAKGIGYFVLQPKVVDGTPCSPDSTSVCOQCCKAGCDRIIDSKKFDKCGVCGGN 720
111 GSTCKKISGSVTSAPGVDIITPTGATNIEVKQNRGSRNNGSFLAIKAADGTVILN 780
112 GSTCKKISGSVTSAPGVDIITPTGATNIEVKQNRGSRNNGSFLAIKAADGTVILN 780
113 GDTYTLSTLEQDIMYKGVVLRYSGSSAALERIRSFSPKLEPLTIQVLTGVGNALRPKIYTY 840
114 GDTYTLSTLEQDIMYKGVVLRYSGSSAALERIRSFSPKLEPLTIQVLTGVGNALRPKIYTY 840
115 FVKKKESFNAIPTFSAWIEEWGECSCSELGWQRRLVECRDINGOPASECAKEVKPAS 900
116 FVKKKESFNAIPTFSAWIEEWGECSCSELGWQRRLVECRDINGOPASECAKEVKPAS 900
117 TRPCADHPCPQWLGEWSSCKTCGKGYKTKSLKLSHDGGVLSHSDCDPLKPKPHIDF 960
118 TRPCADHPCPQWLGEWSSCKTCGKGYKTKSLKLSHDGGVLSHSDCDPLKPKPHIDF 960
119 CTMAECS 967
120 CTMAECS 967

RESULT 5
US-10-755-889-134
; Sequence 134, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 134
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-755-889-134

Query Match          98.8%; Score 5224; DB 16; Length 967;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 955; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MQRVPEGFGRKRLGSDMGNAERAPGSRSGFVPPTLLLLAAALLAVSDALGRPSEDEEL 60
Db 1 MQRVPEGFGRKRLGSDMGNAERAPGSRSGFVPPTLLLLAAALLAVSDALGRPSEDEEL 60
Qy 61 VPELERVPGHGTTRLRHLHAFDQQLDLDVPPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
Db 61 VPELERAPGHGTTRLRHLHAFDQQLDLELRDPSFLAPGFTLQNVGRKSGSETLPETDL 120
Qy 121 AHCFYSGTVNGDPSSAAALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAP 180
```

```
Db 121 AHCFYSGTVNGDPSSAAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAP 180
Qy 181 LQFHLLRRNRQDVGTCGVVDDEPRPTGKATEDEDETEDEGEDEGEPOWSPDPAALQGVG 240
Db 181 LQFHLLRRNRQDVGTCGVVDDEPRPTGKATEDEDETEDEGEDEGEPOWSPDPAALQGVG 240
Qy 241 QPTGTGSIKKRFVSSHRYVETMLVADQSMABFHGSLGKHYLLTLFSAARLYKHPSIRN 300
Db 241 QPTGTGSIKKRFVSSHRYVETMLVADQSMABFHGSLGKHYLLTLFSAARLYKHPSIRN 300
Qy 301 SVSLVVVKILVHDEQKGPEVTSNAALTLRNFCNWQKQHNPPSDRDAEHYDTAILFTROD 360
Db 301 SVSLVVVKILVHDEQKGPEVTSNAALTLRNFCNWQKQHNPPSDRDAEHYDTAILFTROD 360
Qy 361 LCGSOTCDTLGNADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPNHDDAKQACSL 420
Db 361 LCGSOTCDTLGNADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPNHDDAKQACSL 420
Qy 421 NGVNQDSHMMASMLNLDHSPWSPCSGYMITSFLDNGHGECLMDKPNPIQLPGDLPGT 480
Db 421 NGVNQDSHMMASMLNLDHSPWSPCSGYMITSFLDNGHGECLMDKPNPIQLPGDLPGT 480
Qy 481 SYDANRQOQFTFGEDSKHCPDAASTCTLWCTGTSGVLVVCOTKHPHADGTSCGEGKWC 540
Db 481 SYDANRQOQFTFGEDSKHCPDAASTCTLWCTGTSGVLVVCOTKHPHADGTSCGEGKWC 540
Qy 541 INKCVNKNHRKHFDTPPHGSGMWGMPGDCSRTCCTGGGVQYVTRMRECDNPVPKNGSKYCEG 600
Db 541 INKCVNKNHRKHFDTPPHGSGMWGMPGDCSRTCCTGGGVQYVTRMRECDNPVPKNGSKYCEG 600
Qy 601 KRVRYSNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEMIPKYAGVSPKDRCKL 660
Db 601 KRVRYSNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEMIPKYAGVSPKDRCKL 660
Qy 661 ICQAKGIGYFFVLQPKVVDGTPCSDSTSVCCQGVKAGCDRIIDSKKKPKGVCVCGN 720
Db 661 ICQAKGIGYFFVLQPKVVDGTPCSDSTSVCCQGVKAGCDRIIDSKKKPKGVCVCGN 720
Qy 721 GSTCKKISGSVTSAPGYVHDIITPTGATNIEVKORNGSRNNGSFLAIKAADGTYILN 780
Db 721 GSTCKKISGSVTSAPGYVHDIITPTGATNIEVKORNGSRNNGSFLAIKAADGTYILN 780
Qy 781 GDYTLSTLEQDIMYKGVVLYRSGSSAALERIRSFPLKEPLTIQVLTGVGNALRPKIKYTY 840
Db 781 GDYTLSTLEQDIMYKGVVLYRSGSSAALERIRSFPLKEPLTIQVLTGVGNALRPKIKYTY 840
Qy 841 FVKKKESFNAIPTFSAWVIEBWGECSKSELGWQRRLVECRDINGQPASECAKEVKPAS 900
Db 841 FVKKKESFNAIPTFSAWVIEBWGECSKSELGWQRRLVECRDINGQPASECAKEVKPAS 900
Qy 901 TRPCADHPCPQWLGEWSSCKTCGKYKKTSLKCLSHDGGVLSHSDCDPLKKPKHFDIF 960
Db 901 TRPCADHPCPQWLGEWSSCKTCGKYKKTSLKCLSHDGGVLSHSDCDPLKKPKHFDIF 960
```

## RESULT 6

```
US-10-741-600-1603
; Sequence 1603, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 1603
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1603
```

Query Match 98.8%; Score 5224; DB 17; Length 967;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 956; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```
Qy 1 MORAVPEFGRRKLGSDMGNAERAPGSRSGFVPVPTLLLLAAALLAVSDALGRPSSEDEL 60
Db 1 MORAVPEFGRRKLGSDMGNAERAPGSRSGFVPVPTLLLLAAALLAVSDALGRPSSEDEL 60
Qy 61 VYPELERVPGHGTTLRRLHAFDQQLDLVDVPPDSSFLAFGFTLQNVGRKSGSDTLPETDL 120
Db 61 VYPELERAPGHGTTLRRLHAFDQQLDLRLRPPDSSFLAFGFTLQNVGRKSGSETLPETDL 120
Qy 121 AHCFYSGTVNGDPSAAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAP 180
Db 121 AHCFYSGTVNGDPSAAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAP 180
Qy 181 LQFHLLRRNRQDVGTCGVVDDEPRPTGKATEDEDETEDEGEDEGEPOWSPDPAALQGV 240
Db 181 LQFHLLRRNRQDVGTCGVVDDEPRPTGKATEDEDETEDEGEDEGEPOWSPDPAALQGV 240
Qy 241 QPTGTGSIKKRFVSSHRYVETMLVADQSMABFHGSLGKHYLLTLFSAARLYKHPSIRN 300
Db 241 QPTGTGSIKKRFVSSHRYVETMLVADQSMABFHGSLGKHYLLTLFSAARLYKHPSIRN 300
Qy 301 SVSLVVVKILVHDEQKGPEVTSNAALTLRNFCNWQKQHNPPSDRDAEHYDTAILFTROD 360
Db 301 SVSLVVVKILVHDEQKGPEVTSNAALTLRNFCNWQKQHNPPSDRDAEHYDTAILFTROD 360
Qy 361 LCGSOTCDTLGNADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPNHDDAKQACSL 420
Db 361 LCGSOTCDTLGNADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPNHDDAKQACSL 420
Qy 421 NGVNQDSHMMASMLNLDHSPWSPCSGYMITSFLDNGHGECLMDKPNPIQLPGDLPGT 480
Db 421 NGVNQDSHMMASMLNLDHSPWSPCSGYMITSFLDNGHGECLMDKPNPIQLPGDLPGT 480
Qy 481 SYDANRQOQFTFGEDSKHCPDAASTCTLWCTGTSGVLVVCOTKHPHADGTSCGEGKWC 540
Db 481 SYDANRQOQFTFGEDSKHCPDAASTCTLWCTGTSGVLVVCOTKHPHADGTSCGEGKWC 540
Qy 541 INKCVNKNHRKHFDTPPHGSGMWGMPGDCSRTCCTGGGVQYVTRMRECDNPVPKNGSKYCEG 600
Db 541 INKCVNKNHRKHFDTPPHGSGMWGMPGDCSRTCCTGGGVQYVTRMRECDNPVPKNGSKYCEG 600
Qy 601 KRVRYSNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEMIPKYAGVSPKDRCKL 660
Db 601 KRVRYSNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEMIPKYAGVSPKDRCKL 660
Qy 661 ICQAKGIGYFFVLQPKVVDGTPCSDSTSVCCQGVKAGCDRIIDSKKKPKGVCVCGN 720
Db 661 ICQAKGIGYFFVLQPKVVDGTPCSDSTSVCCQGVKAGCDRIIDSKKKPKGVCVCGN 720
Qy 721 GSTCKKISGSVTSAPGYVHDIITPTGATNIEVKORNGSRNNGSFLAIKAADGTYILN 780
Db 721 GSTCKKISGSVTSAPGYVHDIITPTGATNIEVKORNGSRNNGSFLAIKAADGTYILN 780
Qy 781 GDYTLSTLEQDIMYKGVVLYRSGSSAALERIRSFPLKEPLTIQVLTGVGNALRPKIKYTY 840
Db 781 GDYTLSTLEQDIMYKGVVLYRSGSSAALERIRSFPLKEPLTIQVLTGVGNALRPKIKYTY 840
Qy 841 FVKKKESFNAIPTFSAWVIEBWGECSKSELGWQRRLVECRDINGQPASECAKEVKPAS 900
Db 841 FVKKKESFNAIPTFSAWVIEBWGECSKSELGWQRRLVECRDINGQPASECAKEVKPAS 900
Qy 901 TRPCADHPCPQWLGEWSSCKTCGKYKKTSLKCLSHDGGVLSHSDCDPLKKPKHFDIF 960
Db 901 TRPCADHPCPQWLGEWSSCKTCGKYKKTSLKCLSHDGGVLSHSDCDPLKKPKHFDIF 960
```

```
Qy 961 CTMAECS 967
Db 961 CTMAECS 967

RESULT 7
US-10-741-600-1604
; Sequence 1604, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1604
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1604

Query Match 98.8%; Score 5224; DB 17; Length 967;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 956; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MORAVPEGFRKRLGSDMGNAERAPGSRSGFVPPTLLALLAAALLAVSDALGRPSEDEEL 60
Db 1 MORAVPEGFRKRLGSDMGNAERAPGSRSGFVPPTLLALLAAALLAVSDALGRPSEDEEL 60

Qy 61 VPELERVPCHGTTTLRLHAFDQDLDPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
Db 61 VPELERVPCHGTTTLRLHAFDQDLDPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120

Qy 61 VPELERAPGHTTTLRLHAFDQDLDPDSSFLAPGFTLQNVGRKSGSETLPETDL 120
Db 61 VPELERAPGHTTTLRLHAFDQDLDPDSSFLAPGFTLQNVGRKSGSETLPETDL 120

Qy 121 AHCFYSGTVNGDPSAAALSCEGVGAFYLLGAEYFIQPLPAASERLATAAPEKPPAP 180
Db 121 AHCFYSGTVNGDPSAAALSCEGVGAFYLLGAEYFIQPLPAASERLATAAPEKPPAP 180

Qy 121 AHCFYSGTVNGDPSAAALSCEGVGAFYLLGAEYFIQPLPAASERLATAAPEKPPAP 180
Db 121 AHCFYSGTVNGDPSAAALSCEGVGAFYLLGAEYFIQPLPAASERLATAAPEKPPAP 180

Qy 181 LQFHLLRRNQDVGTVCDPSRSCVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
Db 181 LQFHLLRRNQDVGTVCDPSRSCVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420

Qy 181 LQFHLLRRNQDVGTVCDPSRSCVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
Db 181 LQFHLLRRNQDVGTVCDPSRSCVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420

Qy 421 NGVNQDSHMAASMLNLDHSPWPCSGYMITSLDNGHGECLMDKQNPQLPGDLPGT 480
Db 421 NGVNQDSHMAASMLNLDHSPWPCSGYMITSLDNGHGECLMDKQNPQLPGDLPGT 480

Qy 481 SYDANRQCOFTFGEDSKHCPDAASTCTLWCTGTSGGVLVCQTQHPWADGTCGEGKWC 540
Db 481 SYDANRQCOFTFGEDSKHCPDAASTCTLWCTGTSGGVLVCQTQHPWADGTCGEGKWC 540

Qy 541 INKCVNKNHRKHFDTPPHGSGWMGMPGDCSRCTCGGQVQVTMRECDNVPKNGKCEG 600
Db 541 INKCVNKNHRKHFDTPPHGSGWMGMPGDCSRCTCGGQVQVTMRECDNVPKNGKCEG 600

Qy 601 KRVYRSCNLEDPCDNNKGTFRBQCEAHNEFSKASFGSGPAVEMI PKYAGVSPKDRCKL 660
Db 601 KRVYRSCNLEDPCDNNKGTFRBQCEAHNEFSKASFGSGPAVEMI PKYAGVSPKDRCKL 660
```

```
Qy 661 ICQAKGIGYFVLPQKVVDTGTPCSPDSTSVCVQGCQKAGCDRIIDSKKKFDKCGVCGN 720
Db 661 ICQAKGIGYFVLPQKVVDTGTPCSPDSTSVCVQGCQKAGCDRIIDSKKKFDKCGVCGN 720

Qy 721 GSTCKKISGSVTSAPKPGYHDIITITPTGATNIEVKORNQSRNNGSFLAIKAADGTIYN 780
Db 721 GSTCKKISGSVTSAPKPGYHDIITITPTGATNIEVKORNQSRNNGSFLAIKAADGTIYN 780

Qy 781 GDYTLTLEQDQIMYKGVVLYRYSGSSAALERIRSFSPKLEPTIQVLTGVALRPKIKYTY 840
Db 781 GDYTLTLEQDQIMYKGVVLYRYSGSSAALERIRSFSPKLEPTIQVLTGVALRPKIKYTY 840

Qy 841 FVKKKESFNAIPTFSAWVTEWCEKSCSCELGHQORLVECRDINGOPASECAKEVPAS 900
Db 841 FVKKKESFNAIPTFSAWVTEWCEKSCSCELGHQORLVECRDINGOPASECAKEVPAS 900

Qy 901 TRPCADHPCPQWLGEWSSCKTCGKGYKTSKLSKLSHDGVLSHDSCDPLKKPKHFIDF 960
Db 901 TRPCADHPCPQWLGEWSSCKTCGKGYKTSKLSKLSHDGVLSHDSCDPLKKPKHFIDF 960

Qy 961 CTMAECS 967
Db 961 CTMAECS 967

RESULT 8
US-10-923-035-37
; Sequence 37, Application US/10923035
; Publication No. US20050130189A1
; GENERAL INFORMATION:
; APPLICANT: Pasricha, Pankaj
; APPLICANT: Shenoy, Mohan
; APPLICANT: Winston, John
; TITLE OF INVENTION: Compositions and Methods for Treating and Diagnosing
; FILE OF INVENTION: Irritable Bowel Syndrome
; FILE REFERENCE: 9511-136-27
; CURRENT APPLICATION NUMBER: US/10/923,035
; CURRENT FILING DATE: 2004-08-23
; PRIOR APPLICATION NUMBER: US 60/496,716
; PRIOR FILING DATE: 2003-08-21
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-923-035-37

Query Match 98.8%; Score 5224; DB 18; Length 967;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 956; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MORAVPEGFRKRLGSDMGNAERAPGSRSGFVPPTLLALLAAALLAVSDALGRPSEDEEL 60
Db 1 MORAVPEGFRKRLGSDMGNAERAPGSRSGFVPPTLLALLAAALLAVSDALGRPSEDEEL 60

Qy 61 VPELERVPCHGTTTLRLHAFDQDLDPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
Db 61 VPELERVPCHGTTTLRLHAFDQDLDPDSSFLAPGFTLQNVGRKSGSETLPETDL 120

Qy 121 AHCFYSGTVNGDPSAAALSCEGVGAFYLLGAEYFIQPLPAASERLATAAPEKPPAP 180
Db 121 AHCFYSGTVNGDPSAAALSCEGVGAFYLLGAEYFIQPLPAASERLATAAPEKPPAP 180

Qy 181 LQFHLLRRNQDVGTVCDPSRSCVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
Db 181 LQFHLLRRNQDVGTVCDPSRSCVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420

Qy 421 NGVNQDSHMAASMLNLDHSPWPCSGYMITSLDNGHGECLMDKQNPQLPGDLPGT 480
Db 421 NGVNQDSHMAASMLNLDHSPWPCSGYMITSLDNGHGECLMDKQNPQLPGDLPGT 480

Qy 481 SYDANRQCOFTFGEDSKHCPDAASTCTLWCTGTSGGVLVCQTQHPWADGTCGEGKWC 540
Db 481 SYDANRQCOFTFGEDSKHCPDAASTCTLWCTGTSGGVLVCQTQHPWADGTCGEGKWC 540

Qy 541 INKCVNKNHRKHFDTPPHGSGWMGMPGDCSRCTCGGQVQVTMRECDNVPKNGKCEG 600
Db 541 INKCVNKNHRKHFDTPPHGSGWMGMPGDCSRCTCGGQVQVTMRECDNVPKNGKCEG 600

Qy 601 KRVYRSCNLEDPCDNNKGTFRBQCEAHNEFSKASFGSGPAVEMI PKYAGVSPKDRCKL 660
Db 601 KRVYRSCNLEDPCDNNKGTFRBQCEAHNEFSKASFGSGPAVEMI PKYAGVSPKDRCKL 660
```



Db 301 SVSLVVVKILVIHDEQKGEVTSNAALTLRNFCWQKHNPSPDRDAEHYDTAILFTROD 360  
Qy 361 LCGSOTCDTLGWADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKOCASL 420  
Db 361 LCGSOTCDTLGWADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKOCASL 420  
Qy 421 NGVNQDSHMMASMLNLDHSQPSWCSGYMITSFLDNGHGECLMDKPQNPIQLPGDLPGT 480  
Db 421 NGVNQDSHMMASMLNLDHSQPSWCSGYMITSFLDNGHGECLMDKPQNPIQLPGDLPGT 480  
Qy 481 SYDANRQCQFTFGEDSKHCPDAASTCTLWCTGTSGGVLCVOTKHPWADGTSCEGKWC 540  
Db 481 SYDANRQCQFTFGEDSKHCPDAASTCTLWCTGTSGGVLCVOTKHPWADGTSCEGKWC 540  
Qy 541 INKGCVNKHRRKHFDTPPFHSGWGMGPMWGDGCSRTCGGGVQYTMRECDNVPKNGGKYCEG 600  
Db 541 INKGCVNKTRKHFDTPPFHSGWGMGPMWGDGCSRTCGGGVQYTMRECDNVPKNGGKYCEG 600  
Qy 601 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660  
Db 601 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660  
Qy 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSKKKFKCGVCGGN 720  
Db 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSKKKFKCGVCGGN 720  
Qy 721 GSTCKKISGVSATKPGVHDIITPTGATNIEVKQNRQGRNNGSFLAIKAADGTYILN 780  
Db 721 GSTCKKISGVSATKPGVHDIITPTGATNIEVKQNRQGRNNGSFLAIKAADGTYILN 780  
Qy 781 GDYTLSTLEQDIMYKGVVLYRSGSSAALERSFSPLKEPLTIQVLTGVNLRPKIKYTY 840  
Db 781 GDYTLSTLEQDIMYKGVVLYRSGSSAALERSFSPLKEPLTIQVLTGVNLRPKIKYTY 840  
Qy 841 FVKKKKESFNALPTFSAWVIEBWGECSKSCELGWQRRLVECRDINGQPASECAKEVKPAS 900  
Db 841 FVKKKKESFNALPTFSAWVIEBWGECSKSCELGWQRRLVECRDINGQPASECAKEVKPAS 900  
Qy 901 TRPCADHPCPQWLGEWSSCSKTCGKYKTKSLKCLSHDGGVLSHSDCDPLKPKHFIDF 960  
Db 901 TRPCADHPCPQWLGEWSSCSKTCGKYKTKSLKCLSHDGGVLSHSDCDPLKPKHFIDF 960

RESULT 9

US-09-373-658-125  
; Sequence 125, Application US/09373658  
; Publication No. US20030092900A1  
; GENERAL INFORMATION:  
; APPLICANT: Iruela-Arispe, Luisa  
; APPLICANT: Hastings, Gregg A.  
; APPLICANT: Ruben, Steven M.  
; APPLICANT: Jonak, Zdenka L.  
; APPLICANT: Trulli, Stephen H.  
; APPLICANT: Fronwald, James A.  
; APPLICANT: Terrett, Jonathan A.  
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides  
; FILE REFERENCE: 1488.1070006  
; CURRENT APPLICATION NUMBER: US/09/373,658  
; NUMBER OF SEQ ID NOS: 125  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 125  
; LENGTH: 968  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-373-658-125

Query Match 98.8%; Score 5221.5; DB 10; Length 968;

Best Local Similarity 98.9%; Pred. No. 0;  
Matches 957; Conservative 4; Mismatches 6; Indels 1; Gaps 1;  
Qy 1 MORAVPEGFGRKLGSD-MGNAERAPGSRSGPVPPTLLLLAAALAVSDALGRPSEDEE 59  
Db 1 MORAVPEGFGRKLGSDMMGNAERAPGSRSGPVPPTLLLLAAALAVSDALGRPSEDEE 60  
Qy 60 LVVPELERVPVGHGTTTRLRLHAFDQDQDLDDVPDSSFLAPGFTTLQNVGRKSGSDTLPETD 119  
Db 61 LVVPELERAPVGHGTTTRLRLHAFDQDQDLDDVPDSSFLAPGFTTLQNVGRKSGSETLPETD 120  
Qy 120 LAHCFYSYTVNGDPSAAALSILCEGVRCGAFYLLGEAYFIQPLPAASERLATAAPGEKPPA 179  
Db 121 LAHCFYSYTVNGDPSAAALSILCEGVRCGAFYLLGEAYFIQPLPAASERLATAAPGEKPPA 180  
Qy 180 PLOPHLLLRNRQDVGVTGCVVVDDEPRPTGKAETDEDEGTEGEGQWSPQDPALQGV 239  
Db 181 PLOPHLLLRNRQDVGVTGCVVVDDEPRPTGKAETDEDEGTEGEGQWSPQDPALQGV 240  
Qy 240 GQPTGTGSIIRKRFVSSHRYVETMLVADQSMAPFHGSLKHLYLLTFLSVAARLYKHPISIR 299  
Db 241 GQPTGTGSIIRKRFVSSHRYVETMLVADQSMAPFHGSLKHLYLLTFLSVAARLYKHPISIR 300  
Qy 300 NSVSLVVVKILVIHDEQKGEVTSNAALTLRNFCWQKHNPSPDRDAEHYDTAILFTRO 359  
Db 301 NSVSLVVVKILVIHDEQKGEVTSNAALTLRNFCWQKHNPSPDRDAEHYDTAILFTRO 360  
Qy 360 DLCSGOTCDTLGMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKOCAS 419  
Db 361 DLCSGOTCDTLGMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKOCAS 420  
Qy 420 LNGVNQDSHMMASMLNLDHSQPSWCSGYMITSFLDNGHGECLMDKPQNPIQLPGDLPG 479  
Db 421 LNGVNQDSHMMASMLNLDHSQPSWCSGYMITSFLDNGHGECLMDKPQNPIQLPGDLPG 480  
Qy 480 TSYDANRQCQFTFGEDSKHCPDAASTCTLWCTGTSGGVLCVOTKHPWADGTSCEGKW 539  
Db 481 TSYDANRQCQFTFGEDSKHCPDAASTCTLWCTGTSGGVLCVOTKHPWADGTSCEGKW 540  
Qy 540 CINGKCVNKHRRKHFDTPPFHSGWGMGPMWGDGCSRTCGGGVQYTMRECDNVPKNGGKYCE 599  
Db 541 CINGKCVNKHRRKHFDTPPFHSGWGMGPMWGDGCSRTCGGGVQYTMRECDNVPKNGGKYCE 600  
Qy 600 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCK 659  
Db 601 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCK 660  
Qy 660 LIQAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSKKKFKCGVCGG 719  
Db 661 LIQAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSKKKFKCGVCGG 720  
Qy 720 NGSTCKKISGVSATKPGVHDIITPTGATNIEVKQNRQGRNNGSFLAIKAADGTYIL 779  
Db 721 NGSTCKKISGVSATKPGVHDIITPTGATNIEVKQNRQGRNNGSFLAIKAADGTYIL 780  
Qy 780 NGDYTLSTLEQDIMYKGVVLYRSGSSAALERSFSPLKEPLTIQVLTGVNLRPKIKYTY 839  
Db 781 NGDYTLSTLEQDIMYKGVVLYRSGSSAALERSFSPLKEPLTIQVLTGVNLRPKIKYTY 840  
Qy 840 YFVKKKKESFNALPTFSAWVIEBWGECSKSCELGWQRRLVECRDINGQPASECAKEVKPA 899  
Db 841 YFVKKKKESFNALPTFSAWVIEBWGECSKSCELGWQRRLVECRDINGQPASECAKEVKPA 900  
Qy 900 STRPCADHPCPQWLGEWSSCSKTCGKYKTKSLKCLSHDGGVLSHSDCDPLKPKHFID 959  
Db 901 STRPCADHPCPQWLGEWSSCSKTCGKYKTKSLKCLSHDGGVLSHSDCDPLKPKHFID 960  
Qy 960 FCTMAECS 967  
Db 961 FCTMAECS 968

RESULT 10

```

US-09-373-658-2
; Sequence 2, Application US/09373658
; Publication No. US20030092900A1
; GENERAL INFORMATION:
; APPLICANT: Iruela-Arispe, Luisa
; APPLICANT: Hastings, Gregg A.
; APPLICANT: Ruben, Steven M.
; APPLICANT: Jonak, Zdenka L.
; APPLICANT: Trulli, Stephen H.
; APPLICANT: Fronwald, James A.
; APPLICANT: Terrett, Jonathan A.
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides
; FILE REFERENCE: 1488.1070006
; CURRENT APPLICATION NUMBER: US/09/373,658
; CURRENT FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-373-658-2

Query Match      97.3%; Score 5144; DB 10; Length 950;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 940; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 18 MGAERAPGSRSGFVPVPTLLLLAAALLAVSDALGRPSEDEELVPELVRPGHGTTLRL 77
DQ 1 MGAERAPGSRSGFVPVPTLLLLAAALLAVSDALGRPSEDEELVPELVRPGHGTTLRL 60
QY 78 LHAFDQQLDLVPPDSSFLAPGFTLQNVGRKSGSDTLPETDLAHCFYSGTVNGDPSSAA 137
DB 61 LHAFDQQLDLVPPDSSFLAPGFTLQNVGRKSGSETLPETDLAHCFYSGTVNGDPSSAA 120
QY 138 ALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLQFHLLRRNRQGDVGGT 197
DB 121 ALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLQFHLLRRNRQGDVGGT 180
QY 198 CGVVDDEPRPTGKAETDEDEGTEGEGPQSPQDPALQGVQPTGTGSIKKRFVSSH 257
DB 181 CGVVDDEPRPTGKAETDEDEGTEGEGPQSPQDPALQGVQPTGTGSIKKRFVSSH 240
QY 258 RYVETMLVADQSMAEFHGSLGHYLLTLFSAARLYKHPSIRNSVSLVVKILVHDEQK 317
DB 241 RYVETMLVADQSMAEFHGSLGHYLLTLFSAARLYKHPSIRNSVSLVVKILVHDEQK 300
QY 318 GPEVTSNAALTNRNFCNWKQHNPPSDRDAEHYDTAILFTRQDLCSGOTCDTLGMADVGT 360
DB 378 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACASLNGVNDSSHMA SMLSNL 437
DB 361 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACASLNGVNDSSHMA SMLSNL 420
QY 438 DHSQPWSPCSGYMITSFLDNGHCECLMDKQNPQLPGDLPGTSYDANRQCQFTFGEDSK 497
DB 421 DHSQPWSPCSGYMITSFLDNGHCECLMDKQNPQLPGDLPGTSYDANRQCQFTFGEDSK 480
QY 498 HCPDAAASTCSTLWCTGTSGGVVLVQTKHFPWADGTSCEGKWCINGKCNKQNRKHGFDTP 557
DB 481 HCPDAAASTCSTLWCTGTSGGVVLVQTKHFPWADGTSCEGKWCINGKCNKQNRKHGFDTP 540
QY 558 FHSGWGMWPGWDCSRCTCGGVQVYTMRECDNPVKNKGKCEGKRVYRSCNLEDPCDNN 617
DB 541 FHSGWGMWPGWDCSRCTCGGVQVYTMRECDNPVKNKGKCEGKRVYRSCNLEDPCDNN 600
QY 618 KGTFREEQCEAHNBFKASFGSGPAVEWIPKYAGVSPKDRCKLICOAKGIGYFVLQPKV 677
DB 601 KGTFREEQCEAHNBFKASFGSGPAVEWIPKYAGVSPKDRCKLICOAKGIGYFVLQPKV 660
QY 678 VDGTPCSPDSTSVQGVQCVKAGCDRIIDSKKKFKDKCGVCGGNGSTCKKISGVSITSAKPG 737
DB 660 VDGTPCSPDSTSVQGVQCVKAGCDRIIDSKKKFKDKCGVCGGNGSTCKKISGVSITSAKPG 720

```

```

DB 661 VDGTPCSPDSTSVQGVQCVKAGCDRIIDSKKKFKDKCGVCGGNGSTCKKISGVSITSAKPG 720
QY 738 YHDIITPTGATNLEVKORNGSRNNGSFIAIKAADGTYLNGDYTLSTLEQDIMYKGV 797
DB 721 YHDIITPTGATNLEVKORNGSRNNGSFIAIKAADGTYLNGDYTLSTLEQDIMYKGV 780
QY 798 VLRYSGSSAALERIRSFPLKEPLTIQVLTVGNALRPKIKYTVFKKKKESFNAIPTESA 857
DB 781 VLRYSGSSAALERIRSFPLKEPLTIQVLTVGNALRPKIKYTVFKKKKESFNAIPTESA 840
QY 858 WVIWEGECSKSCBLGWQRRLVECRDINGQPASCAKEVKPASTRCPADHPCPQWQLGEW 917
DB 841 WVIWEGECSKSCBLGWQRRLVECRDINGQPASCAKEVKPASTRCPADHPCPQWQLGEW 900
QY 918 SSCSKTCGKYKKTSLKCLSHDGGVLSHSDCDPLKPKKHFIDFCTMAECS 967
DB 901 SSCSKTCGKYKKTSLKCLSHDGGVLSHSDCDPLKPKKHFIDFCTMAECS 950

RESULT 11
US-09-989-687-2
; Sequence 2, Application US/09989687
; Publication No. US2004000249A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides
; FILE REFERENCE: 1488.107000D
; CURRENT APPLICATION NUMBER: US/09/989,687
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-989-687-2

Query Match      97.3%; Score 5144; DB 11; Length 950;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 940; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 18 MGAERAPGSRSGFVPVPTLLLLAAALLAVSDALGRPSEDEELVPELVRPGHGTTLRL 77
DB 1 MGAERAPGSRSGFVPVPTLLLLAAALLAVSDALGRPSEDEELVPELVRPGHGTTLRL 60
QY 78 LHAFDQQLDLVPPDSSFLAPGFTLQNVGRKSGSDTLPETDLAHCFYSGTVNGDPSSAA 137
DB 61 LHAFDQQLDLVPPDSSFLAPGFTLQNVGRKSGSETLPETDLAHCFYSGTVNGDPSSAA 120
QY 138 ALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLQFHLLRRNRQGDVGGT 197
DB 121 ALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLQFHLLRRNRQGDVGGT 180
QY 198 CGVVDDEPRPTGKAETDEDEGTEGEGPQSPQDPALQGVQPTGTGSIKKRFVSSH 257
DB 181 CGVVDDEPRPTGKAETDEDEGTEGEGPQSPQDPALQGVQPTGTGSIKKRFVSSH 240
QY 258 RYVETMLVADQSMAEFHGSLGHYLLTLFSAARLYKHPSIRNSVSLVVKILVHDEQK 317
DB 241 RYVETMLVADQSMAEFHGSLGHYLLTLFSAARLYKHPSIRNSVSLVVKILVHDEQK 300
QY 318 GPEVTSNAALTNRNFCNWKQHNPPSDRDAEHYDTAILFTRQDLCSGOTCDTLGMADVGT 377
DB 301 GPEVTSNAALTNRNFCNWKQHNPPSDRDAEHYDTAILFTRQDLCSGOTCDTLGMADVGT 360
QY 378 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACASLNGVNDSSHMA SMLSNL 437
DB 361 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACASLNGVNDSSHMA SMLSNL 420
QY 438 DHSQPWSPCSGYMITSFLDNGHCECLMDKQNPQLPGDLPGTSYDANRQCQFTFGEDSK 497
DB 421 DHSQPWSPCSGYMITSFLDNGHCECLMDKQNPQLPGDLPGTSYDANRQCQFTFGEDSK 480

```

```
QY 498 HCPDAASTCSTLWCTGTSGVLCVQTKHPFPAWADGTSCEGKWCINGKCVNKNRKHFDTP 557
|
|
|
Db 481 HCPDAASTCSTLWCTGTSGVLCVQTKHPFPAWADGTSCEGKWCINGKCVNKNRKHFDTP 540
|
|
|
QY 558 FHGSGMGPWGDGCSRTCGGGVQYTMRECDNVPKNGGKYCEGKVRVRSCHLEDCPDNN 617
|
|
|
Db 541 FHGSGMGPWGDGCSRTCGGGVQYTMRECDNVPKNGGKYCEGKVRVRSCHLEDCPDNN 600
|
|
|
QY 618 GXTFREEOCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLI COAKGIGYFFVLQPKV 677
|
|
|
Db 601 GXTFREEOCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLI COAKGIGYFFVLQPKV 660
|
|
|
QY 678 VDGTPCSPDSTSVQVQGCVKAGCDRIIDSXKFKDKCGVCGNGSTCKKISGSVTSAPKG 737
|
|
|
Db 661 VDGTPCSPDSTSVQVQGCVKAGCDRIIDSXKFKDKCGVCGNGSTCKKISGSVTSAPKG 720
|
|
|
QY 738 YHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYYILNGDYTLSTLEQDIMYKGV 797
|
|
|
Db 721 YHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYYILNGDYTLSTLEQDIMYKGV 780
|
|
|
QY 798 VLRYSGSSAALERIRSFSPKLEPLTIQVLTGNALRPKIKYTYFVKKKKESFNAIPTFSA 857
|
|
|
Db 781 VLRYSGSSAALERIRSFSPKLEPLTIQVLTGNALRPKIKYTYFVKKKKESFNAIPTFSA 840
|
|
|
QY 858 WVEEWGECSSCELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHPCPQWLGEW 917
|
|
|
Db 841 WVEEWGECSSCELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHPCPQWLGEW 900
|
|
|
QY 918 SSCSKTCGKYKKTSLKCLSHDGGVLSHSDCDPLKKPKHFIDFCTMAECS 967
|
|
|
Db 901 SSCSKTCGKYKKTSLKCLSHDGGVLSHSDCDPLKKPKHFIDFCTMAECS 950
|
|
|
```

## RESULT 12

```
US-10-667-281-2
; Sequence 2, Application US/10667281
; Publication No. US20050100916A1
; GENERAL INFORMATION:
; APPLICANT: Klonowski, Paul
; APPLICANT: Allard, John
; APPLICANT: Heller, Renu
; APPLICANT: Van Wart, Harold
; TITLE OF INVENTION: Human Aggreganase and Nucleic Acid
; TITLE OF INVENTION: Compositions Encoding the Same
; FILE REFERENCE: ROCH-002
; CURRENT APPLICATION NUMBER: US/10/667,281
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: US/09/568,559
; PRIOR FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: 60/133,343
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 949
; TYPE: PRT
; ORGANISM: human
US-10-667-281-2
```

```
Query Match 97.2%; Score 5140; DB 17; Length 949;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 939; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 19 GNAERAPGSRSGPVTLLLLAALLAVSDALGRPSEDEELVPELERVPGHGTTRL 78
|
|
|
Db 1 GNAERAPGSRSGPVTLLLLAALLAVSDALGRPSEDEELVPELERVPGHGTTRL 60
|
|
|
QY 79 HAFDQOLDLVDPPDGSFLAPGFTLQVGRKSGSDTLPETDLAHCFSYGTVNGDPSAAA 138
|
|
|
Db 61 HAFDQOLDLVDPPDGSFLAPGFTLQVGRKSGSDTLPETDLAHCFSYGTVNGDPSAAA 120
|
|
|
QY 139 LSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAPLQPHLLRRNRQGDVGTC 198
|
|
|
```

```
Db 121 LSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAPLQPHLLRRNRQGDVGTC 180
|
|
|
QY 199 GVVDDEPRPTKAEATEDDEDETEGEDEGPQWSPDPAQGVQOPTGTGSIKKRKFVSSH 258
|
|
|
Db 181 GVVDDEPRPTKAEATEDDEDETEGEDEGAQWSPDPAQGVQOPTGTGSIKKRKFVSSH 240
|
|
|
QY 259 YVETMLVADQSMABPHGSLKHLYLTLFSAARLYKHPSIRNSVSLVVKILVHDEQK 318
|
|
|
Db 241 YVETMLVADQSMABPHGSLKHLYLTLFSAARLYKHPSIRNSVSLVVKILVHDEQK 300
|
|
|
QY 319 PEVTSNAALTNRNENWQKHNPPSDRDAEHYDTAILFTRODLCSQTCDTLGMADVTV 378
|
|
|
Db 301 PEVTSNAALTNRNENWQKHNPPSDRDAEHYDTAILFTRODLCSQTCDTLGMADVTV 360
|
|
|
QY 379 CDPSSRSCSVIEDDGLQAAFTTAHELGHVFNPNPHDDAKOCASLNGVQDSHMMASMLSLND 438
|
|
|
Db 361 CDPSSRSCSVIEDDGLQAAFTTAHELGHVFNPNPHDDAKOCASLNGVQDSHMMASMLSLND 420
|
|
|
QY 439 HSQWSPSCGYMITSFLDNHGECLMDKPQNPLOLPGDLPGTSYDANRQCOFTFGEDESKH 498
|
|
|
Db 421 HSQWSPSCGYMITSFLDNHGECLMDKPQNPLOLPGDLPGTSYDANRQCOFTFGEDESKH 480
|
|
|
QY 499 CPDAASTCSTLWCTGTSGVLCVQTKHPFPAWADGTSCEGKWCINGKCVNKNRKHFDTP 558
|
|
|
Db 481 CPDAASTCSTLWCTGTSGVLCVQTKHPFPAWADGTSCEGKWCINGKCVNKNRKHFDTP 540
|
|
|
QY 559 HSGWMGPWGDGCSRTCGGGVQYTMRECDNVPKNGGKYCEGKVRVRSCHLEDCPDNN 618
|
|
|
Db 541 HSGWMGPWGDGCSRTCGGGVQYTMRECDNVPKNGGKYCEGKVRVRSCHLEDCPDNN 600
|
|
|
QY 619 KTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLI COAKGIGYFFVLQPKV 678
|
|
|
Db 601 KTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLI COAKGIGYFFVLQPKV 660
|
|
|
QY 679 DGTFCSPDSTSVQVQGCVKAGCDRIIDSXKFKDKCGVCGNGSTCKKISGSVTSAPKG 738
|
|
|
Db 661 DGTFCSPDSTSVQVQGCVKAGCDRIIDSXKFKDKCGVCGNGSTCKKISGSVTSAPKG 720
|
|
|
QY 739 HDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYYILNGDYTLSTLEQDIMYKGV 798
|
|
|
Db 721 HDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYYILNGDYTLSTLEQDIMYKGV 780
|
|
|
QY 799 LRYSGSSAALERIRSFSPKLEPLTIQVLTGNALRPKIKYTYFVKKKKESFNAIPTFSA 858
|
|
|
Db 781 LRYSGSSAALERIRSFSPKLEPLTIQVLTGNALRPKIKYTYFVKKKKESFNAIPTFSA 840
|
|
|
QY 859 VIEEWGECSSCELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHPCPQWLGEW 918
|
|
|
Db 841 VIEEWGECSSCELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHPCPQWLGEW 900
|
|
|
QY 919 SSCSKTCGKYKKTSLKCLSHDGGVLSHSDCDPLKKPKHFIDFCTMAECS 967
|
|
|
Db 901 SSCSKTCGKYKKTSLKCLSHDGGVLSHSDCDPLKKPKHFIDFCTMAECS 949
|
|
|
```

## RESULT 13

```
US-10-741-600-1605
; Sequence 1605, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1605
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-10-741-600-1605

```

Query Match          97.1%; Score 5136; DB 17; Length 950;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 939; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 18 MGNAERAPGSRSGFVPPTLLALLAAALAVSDALGRPSEDEELVPELERVPHGHTTTLR 77
Db 1 MGNAERAPGSRSGFVPPTLLALLAAALAVSDALGRPSEDEELVPELERVPHGHTTTLR 60

QY 78 LHAFOQDLDDVPDPSSFLAPGFTLQNVGRKSGSDTLPETDLAHCFFVSGTVNGDPSSAA 137
Db 61 LHAFOQDLLELRDPSSFLAPGFTLQNVGRKSGSETPLPETDLAHCFFVSGTVNGDPSSAA 120

QY 138 ALSCEGVGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLOPHLLRRNQDVGVT 197
Db 121 ALSCEGVGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLOPHLLRRNQDVGVT 180

QY 198 CGVVDDEPRPTGKAETDEDEGTEGEDEGQWSPDPALQGVGQPTGTGSIKRRFVSSH 257
Db 181 CGVVDDEPRPTGKAETDEDEGTEGEDEGQWSPDPALQGVGQPTGTGSIKRRFVSSH 240

QY 258 RYVETMLVADQSMABFHGSGGLKHYLLTLFSVAARLYKHPISIRNSVSLVVKILVIHDEQK 317
Db 241 RYVETMLVADQSMABFHGSGGLKHYLLTLFSVAARLYKHPISIRNSVSLVVKILVIHDEQK 300

QY 318 GPEVTSNAALTLRNFCNWQKHNPSPDRDAEHYDTAILFTRODLCSGQTCDTLGMADVGT 377
Db 301 GPEVTSNAALTLRNFCNWQKHNPSPDRDAEHYDTAILFTRODLCSGQTCDTLGMADVGT 360

QY 378 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQCASLNGVNDSHMMASMLSNL 437
Db 361 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQCASLNGVNDSHMMASMLSNL 420

QY 438 DHSQWSPSCGYWITSTFLDNGHGECLMDKPNQPIQLPGDLPGTSYDANRQCFTEGDSK 497
Db 421 DHSQWSPSCGYWITSTFLDNGHGECLMDKPNQPIQLPGDLPGTSYDANRQCFTEGDSK 480

QY 498 HCPDAASTCTLMWCTGTSGGVLVCOTKHFPWADGTSCEGKWCINGKCNVNRHKGHFTDP 557
Db 481 HCPDAASTCTLMWCTGTSGGVLVCOTKHFPWADGTSCEGKWCINGKCNVNRHKGHFTDP 540

QY 558 FHGSGWGMGPWGDCSRTCGGVQYTMRECDNPVPKNGGKCEGKRVYRSCNLEDCPDNN 617
Db 541 FHGSGWGMGPWGDCSRTCGGVQYTMRECDNPVPKNGGKCEGKRVYRSCNLEDCPDNN 600

QY 618 GKTFRBEOCEAHNEFSAKSFSGGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKV 677
Db 601 GKTFRBEOCEAHNEFSAKSFSGGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKV 660

QY 678 VDGTPCSPDSTSVCGQCVKAGCDRIIDSKKKFDKCGVCGNGSTCKKISGSVTSAPKG 737
Db 661 VDGTPCSPDSTSVCGQCVKAGCDRIIDSKKKFDKCGVCGNGSTCKKISGSVTSAPKG 720

QY 738 YHDIITITPTGATNIEVKORNGSRNNGSFLAIKAADGTYILNGDYTLSTLEQDIMYKGV 797
Db 721 YHDIITITPTGATNIEVKORNGSRNNGSFLAIKAADGTYILNGDYTLSTLEQDIMYKGV 780

QY 798 VLRYSGSSAALERIRSFPLKEPLTIQVLTVGNALRPKIKTYTFVKKKESFNAIPTPSA 857
Db 781 VLRYSGSSAALERIRSFPLKEPLTIQVLTVGNALRPKIKTYTFVKKKESFNAIPTPSA 840

QY 858 WVEBWGCSKSELGWORRLVECHDINGQAPASECAKEVKPASTRPCADHPCPQWOLGEW 917
Db 841 WVEBWGCSKSELGWORRLVECHDINGQAPASECAKEVKPASTRPCADHPCPQWOLGEW 900

QY 918 SSCSKTCGKYKTSKLSCLSHDGGVLSHSDCDPLKKPKHFIDFCTMAECS 967
Db 901 SSCSKTCGKYKTSKLSCLSHDGGVLSHSDCDPLKKPKHFIDFCTMAECS 950

```

RESULT 14

US-10-973-858-14

```

; Sequence 14, Application US/10973858
; Publication No. US20050176030A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Li
; APPLICANT: Gonzalez-Zulueta, Mirella
; APPLICANT: Ye, Shiming
; APPLICANT: Urfer, Roman
; APPLICANT: Nickolich, Karoly
; TITLE OF INVENTION: Regulated Nucleic Acids in Pathogenesis of Alzheimer's Disease
; FILE REFERENCE: AGYT-047
; CURRENT APPLICATION NUMBER: US/10/973,858
; CURRENT FILING DATE: 2004-10-25
; PRIOR APPLICATION NUMBER: 60/515,562
; PRIOR FILING DATE: 2003-10-28
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 14
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-973-858-14

```

```

Query Match          97.1%; Score 5136; DB 18; Length 950;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 939; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 18 MGNAERAPGSRSGFVPPTLLALLAAALAVSDALGRPSEDEELVPELERVPHGHTTTLR 77
Db 1 MGNAERAPGSRSGFVPPTLLALLAAALAVSDALGRPSEDEELVPELERVPHGHTTTLR 60

QY 78 LHAFOQDLDDVPDPSSFLAPGFTLQNVGRKSGSDTLPETDLAHCFFVSGTVNGDPSSAA 137
Db 61 LHAFOQDLLELRDPSSFLAPGFTLQNVGRKSGSETPLPETDLAHCFFVSGTVNGDPSSAA 120

QY 138 ALSCEGVGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLOPHLLRRNQDVGVT 197
Db 121 ALSCEGVGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLOPHLLRRNQDVGVT 180

QY 198 CGVVDDEPRPTGKAETDEDEGTEGEDEGQWSPDPALQGVGQPTGTGSIKRRFVSSH 257
Db 181 CGVVDDEPRPTGKAETDEDEGTEGEDEGQWSPDPALQGVGQPTGTGSIKRRFVSSH 240

QY 258 RYVETMLVADQSMABFHGSGGLKHYLLTLFSVAARLYKHPISIRNSVSLVVKILVIHDEQK 317
Db 241 RYVETMLVADQSMABFHGSGGLKHYLLTLFSVAARLYKHPISIRNSVSLVVKILVIHDEQK 300

QY 318 GPEVTSNAALTLRNFCNWQKHNPSPDRDAEHYDTAILFTRODLCSGQTCDTLGMADVGT 377
Db 301 GPEVTSNAALTLRNFCNWQKHNPSPDRDAEHYDTAILFTRODLCSGQTCDTLGMADVGT 360

QY 378 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQCASLNGVNDSHMMASMLSNL 437
Db 361 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQCASLNGVNDSHMMASMLSNL 420

QY 438 DHSQWSPSCGYWITSTFLDNGHGECLMDKPNQPIQLPGDLPGTSYDANRQCFTEGDSK 497
Db 421 DHSQWSPSCGYWITSTFLDNGHGECLMDKPNQPIQLPGDLPGTSYDANRQCFTEGDSK 480

QY 498 HCPDAASTCTLMWCTGTSGGVLVCOTKHFPWADGTSCEGKWCINGKCNVNRHKGHFTDP 557
Db 481 HCPDAASTCTLMWCTGTSGGVLVCOTKHFPWADGTSCEGKWCINGKCNVNRHKGHFTDP 540

QY 558 FHGSGWGMGPWGDCSRTCGGVQYTMRECDNPVPKNGGKCEGKRVYRSCNLEDCPDNN 617
Db 541 FHGSGWGMGPWGDCSRTCGGVQYTMRECDNPVPKNGGKCEGKRVYRSCNLEDCPDNN 600

QY 618 GKTFRBEOCEAHNEFSAKSFSGGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKV 677
Db 601 GKTFRBEOCEAHNEFSAKSFSGGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKV 660

QY 678 VDGTPCSPDSTSVCGQCVKAGCDRIIDSKKKFDKCGVCGNGSTCKKISGSVTSAPKG 737
Db 661 VDGTPCSPDSTSVCGQCVKAGCDRIIDSKKKFDKCGVCGNGSTCKKISGSVTSAPKG 720

```

QY 738 YHDIITPTGATNIEVKQNRGRNNGSFLAIKAADGTYYILNGDYTLSTLSEQDIMYKGV 797  
Db 721 YHDIITPTGATNIEVKQNRGRNNGSFLAIKAADGTYYILNGDYTLSTLSEQDIMYKGV 780  
QY 798 VLRYGSSAALRIIRSFSPKPLTIQVLTGVNALARPKIKYTFYFKKKKESFNAIPTFSA 857  
Db 781 VLRYGSSAALRIIRSFSPKPLTIQVLTGVNALARPKIKYTFYFKKKKESFNAIPTFSA 840  
QY 858 WYIEWGECSKCELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHPQCPQOLGEW 917  
Db 841 WYIEWGECSKCELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHPQCPQOLGEW 900  
QY 918 SSCSKTCGKGYKTKSLKCLSHDGGVLSHSDCLPKPKPHFIDFCTMAECS 967  
Db 901 SSCSKTCGKGYKTKSLKCLSHDGGVLSHSDCLPKPKPHFIDFCTMAECS 950  
RESULT 15  
US-09-741-151-4  
; Sequence 4, Application US/09741151  
; Publication No. US20020086400A1  
; GENERAL INFORMATION:  
; APPLICANT: ZHU, Shaoping et al  
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: CL001005  
; CURRENT APPLICATION NUMBER: US/09/741,151  
; CURRENT FILING DATE: 2000-12-08  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 931  
; TYPE: PRT  
; ORGANISM: Human  
US-09-741-151-4  
Query Match 95.3%; Score 5036; DB 9; Length 931;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 920; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
QY 36 LLLLLAALLAVSDALGRSEDEELVPELERVPGHGTTRLRHLHAFDQQLDLDPDSSF 95  
Db 1 LLLLLAALLAVSDALGRSEDEELVPELERVPGHGTTRLRHLHAFDQQLDLDPDSSF 60  
QY 96 LAPGFTLQNVGRKSGDTPLETDLAHCYFVTVNGDPSSAALSICGVGRGAFYLLGEA 155  
Db 61 LAPGFTLQNVGRKSGSETPLETDLAHCYFVTVNGDPSSAALSICGVGRGAFYLLGEA 120  
QY 156 YFIQPLPAASERLATAAPGKEPPAPLQFHLLRRNRQGVGGTCGVVDDDEPRPTGKAETED 215  
Db 121 YFIQPLPAASERLATAAPGKEPPAPLQFHLLRRNRQGVGGTCGVVDDDEPRPTGKAETED 180  
QY 216 EDEGTEGEDEGPQSPQDPALQGVGQPTGTGIRKRFVSSHRYVETMLVADQSMAEFHG 275  
Db 181 EDEGTEGEDEGAQSPQDPALQGVGQPTGTGIRKRFVSSHRYVETMLVADQSMAEFHG 240  
QY 276 SGLKHVLLTLFSAARLYKHPISVSLVVKILVHDEQKGPVTSNAALTFRNFCNW 335  
Db 241 SGLKHVLLTLFSAARLYKHPISVSLVVKILVHDEQKGPVTSNAALTFRNFCNW 300  
QY 336 QKQHNPPSDRDAEHVDYDTALFTRQDLCSGQTCDTLGMADVGTCDPSRSCSVIEDDGLQA 395  
Db 301 QKQHNPPSDRDAEHVDYDTALFTRQDLCSGQTCDTLGMADVGTCDPSRSCSVIEDDGLQA 360  
QY 396 APTTAHELGHVFNPHDDAKQACASLNGVNDSHMMASMLSNLDHQPWSPSCGYMITSTFL 455  
Db 361 APTTAHELGHVFNPHDDAKQACASLNGVNDSHMMASMLSNLDHQPWSPSCGYMITSTFL 420  
QY 456 DNGHGECLMDKQNPQLPGDLPGTYSYDANROQCTFGEDSKHCPDAASTCTLWCTGTS 515  
Db 421 DNGHGECLMDKQNPQLPGDLPGTYSYDANROQCTFGEDSKHCPDAASTCTLWCTGTS 480

QY 516 GGVLVQCQTKHFPWADGTSCEGKWCINGKCVNKNRKHFDTPFHGSMGMWGPWGDCSRTC 575  
Db 481 GGVLVQCQTKHFPWADGTSCEGKWCINGKCVNKNRKHFDTPFHGSMGMWGPWGDCSRTC 540  
QY 576 GGGVQYTMRECDNPVPKNGGKYCEGKRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKA 635  
Db 541 GGGVQYTMRECDNPVPKNGGKYCEGKRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKA 600  
QY 636 SFGSGPAVEMIPKVAGVSPKDRCKLI COAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQGG 695  
Db 601 SFGSGPAVEMIPKVAGVSPKDRCKLI COAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQGG 660  
QY 696 CVKAGCDRIIDSXXXKFKCGVCGNGSTCKKISGSVTSAKPGYHDIITPTGATNIEVKQ 755  
Db 661 CVKAGCDRIIDSXXXKFKCGVCGNGSTCKKISGSVTSAKPGYHDIITPTGATNIEVKQ 720  
QY 756 RNQGRNNGSFLAIKAADGTYYILNGDYTLSTLSEQDIMYKGVLSGSSAALERIRSF 815  
Db 721 RNQGRNNGSFLAIKAADGTYYILNGDYTLSTLSEQDIMYKGVLSGSSAALERIRSF 780  
QY 816 PLKEPLTIQVLTGVNALARPKIKYTFYFKKKKESFNAIPTFSAWIEWGECSKCELGWQ 875  
Db 781 PLKEPLTIQVLTGVNALARPKIKYTFYFKKKKESFNAIPTFSAWIEWGECSKCELGWQ 840  
QY 876 RRLVECRDINGQPASECAKEVKPASTRPCADHPQCPQOLGEWSSCSKTCGKGYKTKSLKC 935  
Db 841 RRLVECRDINGQPASECAKEVKPASTRPCADHPQCPQOLGEWSSCSKTCGKGYKTKSLKC 900  
QY 936 LSHDGGVLSHSDCLPKPKPHFIDFCTMAEC 966  
Db 901 LSHDGGVLSHSDCLPKPKPHFIDFCTMAEC 931

Search completed: August 22, 2005, 10:23:20  
Job time : 171 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2005, 10:09:09 ; Search time 47 Seconds  
(without alignments)  
1979.609 Million cell updates/sec

Title: US-09-989-687-126

Perfect score: 5287  
Sequence: 1 MRAVPEGFGRKLGSDMGN.....CDPLKKPKHFIDFCTMAECS 967

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 9621673 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4264.5	80.7	951	2 T00017	Gene ADAMTS-1 prot
2	3056	57.8	550	2 T47158	hypothetical prote
3	2129	40.3	837	2 T00355	hypothetical prote
4	1598	30.2	2165	2 T21371	hypothetical prote
5	1096	20.7	1205	2 T18517	procollagen N-endo
6	645	12.2	1558	2 C89114	protein C37C3.6a [
7	645	12.2	2167	2 T34395	hypothetical prote
8	614.5	11.6	1444	2 T18856	angiogenesis inhib
9	466	8.8	860	2 T18892	hypothetical prote
10	418.5	7.9	1059	2 T2545	hypothetical prote
11	402	7.6	957	2 T15976	hypothetical prote
12	391	7.4	951	2 T00260	hypothetical prote
13	378.5	7.2	903	2 S60257	metrin alpha - mo
14	347	6.6	549	2 S48169	metalloproteinase
15	343	6.5	571	2 S24789	Jararagin C.precu
16	343	6.5	617	2 S48160	metalloproteinase
17	340.5	6.4	610	2 JC8056	halyase - Glyoydiu
18	326.5	6.2	610	2 JC7530	vascular apoptosis
19	326.5	6.2	826	2 A60385	monocyte surface a
20	325.5	6.2	609	2 S55270	catrocollastatin p
21	321	6.1	655	2 JC7850	disintegrin and me
22	320.5	6.1	481	2 JC4342	fibrinolytic prote
23	308.5	5.8	481	2 S43125	trimucin precursor
24	304.5	5.8	407	2 S63260	metalloproteinase
25	304	5.7	411	1 HY5NEA	fibrolase (EC 3.4.
26	302	5.7	616	2 A55796	ecarin precursor -
27	301	5.7	478	2 A43296	atrolysin B (EC 3.
28	299	5.7	478	2 JQ1301	hemorrhagic protei
29	296.5	5.6	478	2 JC4880	fibrinolytic metal

ALIGNMENTS

RESULT 1

T00017

Gene ADAMTS-1 protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 15-Mar-2004

C:Accession: T00017

R:Kuno, K.; Lizasa, H.; Ohno, S.; Matsushima, K.

Genomics 46, 466-471, 1997

A>Title: The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1 gene

A:Reference number: Z14055; MUID:98110583; PMID:9441751

A:Accession: T00017

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-951 <KUN>

A:Cross-references: EMBL:AB001735; NID:g2809056; PIDN:BAA24501.1; PID:g2809057

A:Experimental source: strain 129SVJ

C:Genetics:

A:Gene: ADAMTS-1

A:Introns: 228/1; 343/3; 388/1; 444/1; 539/3; 602/1; 660/3; 719/2

F:542-598/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 80.7%; Score 4264.5; DB 2; Length 951;

Best Local Similarity 81.6%; Pred. No. 2.6e-278;

Matches 782; Conservative 56; Mismatches 105; Indels 15; Gaps 5;

Qy	18	MGNARAPGSRSGFVPVTLILLAAA---LLAVSDALGRPSEDEELVPELERYVGHG--	72
Db	1	MGDVQRAARSRSLSAHMLLLLLASITMLLCARGAHRPTDEDEELVLSLERAPGHDS	60
Qy	73	TTRLRLHAFDOQLDLDPDSSFLAPGFTLQNVGRKSGSDTFL--PETDLAHCYSGTVN	130
Db	61	TTRLRLDAFGQQLHLKLPDSDGFLAPGFTLOTVGRSPGSEAHLDPTGDLAHCYSGTVN	120
Qy	131	GDPSAAALSICEGVGRGAFYLLGEAYFTQPLP-AASERLATAAPEKPPAPLQPHLLARN	189
Db	121	GDPSAAALSICEGVGRGAFYLLQGEFFIQAPGAVATERLAPAVEESSARPQPHILARR	180
Qy	190	RQGDVGTCGVVDPEPRPTGKAETDEDEGTEGEDEGPQSPQDPALQGVQPTGSGIR	249
Db	181	RRGGGAKCGVMDDETLP-----SDSRPESQNTQNPVRDPTPDQAKGSGPGSGIR	233
Qy	250	KKRFVSSHRYVETMLVADQSMAEFHGSLKHYLLTLFSVAARLYKHPISIRNSVLVVVKI	309
Db	234	KKRFVSSPRYVETMLVADQSMADFHGSLKHYLLTLFSVAARFYKHPISIRNSISLVVVKI	293
Qy	310	LVIHDEQKGPVTSNAALTTLRNFNCNQKQHPSPDRDDEHYDTAILFTRQDLCSGQCDT	369
Db	294	LVIVEEQKGPVTSNAALTTLRNFNCNQKQHPSPDRDDEHYDTAILFTRQDLCSGHTCDT	353
Qy	370	LGMDVGVTCVDPSPSCSVIEDDGLQAFTTAHELGHVFNMPHDDAKKQACASLNGVNDSHM	429
Db	354	LGMDVGVTCVDPSPSCSVIEDDGLQAFTTAHELGHVFNMPHDDAKKHCASLNGVTGDSHL	413

QY 430 MASMLSNLDHSPWSPCSGYMLTSLFDNGHGCECLMDKPNPIQLPGDLPGTSYDANROCO 489  
 DB 414 MASMLSSLDHSPWSPCSAYMTSLFDNGHGCECLMDKPNPIKLPDSLDPGTLYDANROCO 473  
 QY 490 FTFGSDSKHCPDAASTCTSLWCTGTSGGVLVCQTKHFPWADGTSCEGKWCINGKCVNKN 549  
 DB 474 FTFGESKHCPDAASTCTTLWCTGTSGGLLVCQTKHFPWADGTSCEGKWCVSGKCVNKT 533  
 QY 550 HKKHFDTPPHGSGMWGPGWDCSRTCCTGGGVQVYTMRECDNVPVNGGKYCEGKRVYRSCN 609  
 DB 534 DMKHEPATPVHSGWPGWPGWDCSRTCCTGGGVQVYTMRECDNVPVNGGKYCEGKRVYRSCN 593  
 QY 610 LEDCPDNNKGTFRFQCEAHNEFSKASFGSPAVEWIPKYAGVSPKDRCKLICQAKGIGY 669  
 DB 594 IEDCPDNNKGTFRFQCEAHNEFSKASFGNEFTVEWTPKYAGVSPKDRCKLTCEANGIGY 653  
 QY 670 FVLQPKVVDGTPPCSPDSTSVCCQCVKAGCDRIIDSKKKFDKCGVCGGNGSTCKKISG 729  
 DB 654 FVLQPKVVDGTPPCSPDSTSVCCQCVKAGCDRIIDSKKKFDKCGVCGGNGSTCKKMSG 713  
 QY 730 SVTSAPKPGVHDIITPTGATNIEVKQNRGRNNGSFLAIKAADGTYILNGDYTLSTLE 789  
 DB 714 IVTSTRPGVHDIITIPAGATNIEVKHNRQGRNNGSFLAIRAADGTYILNGNFTLSTLE 773  
 QY 790 ODIMYKGVVLYRSGSSAALERIRSFPLKEPTIOVLTVGNALRPKIYTYFVKKKESF 849  
 DB 774 QDLTYKGTVLYRSGSSAALERIRSFPLKEPTIOVLTVGNALRPKIYTYFVKKKTESF 833  
 QY 850 NAIPTFSAWVIEBWGECSCBLGQRRRLVECRDINGOPASECAKEVPASTRPCADHPC 909  
 DB 834 NAIPTFSEWVIEBWGECSTCGSGWQRRVQCRDINGHPASECAKEVPASTRPCADLPC 893  
 QY 910 POWQJGEWSSCKTCGKGYKTSLKCLSHDGGVLSHSDCDPLKKPHFIDFCTMAECS 967  
 DB 894 PHWQVGDWSPCKTCGKGYKRTKLCVSHDGGVLSNESCDPLKKPHYIDFCTLTQCS 951

RESULT 2  
 T47158  
 hypothetical protein DKFp762C110.1 - human (fragment)  
 C:Species: Homo sapiens (man)  
 C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
 C:Accession: T47158  
 R:Blum, H.; Baurisachs, S.; Mewes, H.W.; Weil, B.; Wiemann, S.  
 submitted to the Protein Sequence Database, March 2000  
 A:Reference number: Z24379  
 A:Accession: T47158  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-550 <AAA>  
 A:Cross-references: UNIPROT:O9UHI8; EMBL:AL162080  
 A:Experimental source: adult melanoma (Mewo cell line); clone DKFzp762C110  
 C:Genetics:  
 A>Note: DKFzp762C110.1

Query Match 57.8%; Score 3056; DB 2; Length 550;  
 Best Local Similarity 99.1%; Pred. No. 1.9e-197;  
 Matches 545; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 418 ASLNGVNDSHMWSMLNLDHSQWSPCSGYMTSFLDNGHGCECLMDKPNPIQLPGDL 477  
 DB 1 ASLNGVNDSHMWSMLNLDHSQWSPCSAYMTSFLDNGHGCECLMDKPNPIQLPGDL 60  
 QY 478 PGTSDANRQCOFTFGEDSKHCPDAASTCTSLWCTGTSGGVLVCQTKHFPWADGTSCEG 537  
 DB 61 PGTSDANRQCOFTFGEDSKHCPDAASTCTSLWCTGTSGGVLVCQTKHFPWADGTSCEG 120  
 QY 538 KWCINGKCVNKHGKHFDPFHGSGMWGPGWDCSRTCCTGGGVQVYTMRECDNVPVNGGKY 597  
 DB 121 KWCINGKCVNKTDRKHFDTPFHGSGMWGPGWDCSRTCCTGGGVQVYTMRECDNVPVNGGKY 180  
 QY 598 CEGKRVYRSCNLEDCPDNNKGTFRFQCEAHNEFSKASFGSPAVEWIPKYAGVSPKDR 657

DB 181 CEGKRVYRSCNLEDCPDNNKGTFRFQCEAHNEFSKASFGSPAVEWIPKYAGVSPKDR 240  
 QY 658 CKLICQAKGIGYFPVLQPKVVDGTPPCSPDSTSVCCQCVKAGCDRIIDSKKKFDKCGVC 717  
 DB 241 CKLICQAKGIGYFPVLQPKVVDGTPPCSPDSTSVCCQCVKAGCDRIIDSKKKFDKCGVC 300  
 QY 718 GGNGSTCKKISGTSVTSAPKPGVHDIITPTGATNIEVKQNRGRNNGSFLAIKAADGTY 777  
 DB 301 GGNGSTCKKISGTSVTSAPKPGVHDIITPTGATNIEVKQNRGRNNGSFLAIKAADGTY 360  
 QY 778 ILNGDYTLSTLEQDIMYKGVVLYRSGSSAALERIRSFPLKEPTIOVLTVGNALRPKIK 837  
 DB 361 ILNGDYTLSTLEQDIMYKGVVLYRSGSSAALERIRSFPLKEPTIOVLTVGNALRPKIK 420  
 QY 838 YTFVKKKESFNAIPTFSAWVIEBWGECSCBLGQRRRLVECRDINGOPASECAKEV 897  
 DB 421 YTFVKKKESFNAIPTFSAWVIEBWGECSCBLGQRRRLVECRDINGOPASECAKEV 480  
 QY 898 PASTRPCADHPCPOWQJGEWSSCKTCGKGYKTSLKCLSHDGGVLSHSDCDPLKKPHF 957  
 DB 481 PASTRPCADHPCPOWQJGEWSSCKTCGKGYKRSKCLSHDGGVLSHSDCDPLKKPHF 540  
 QY 958 IDFCTMAECS 967  
 DB 541 IDFCTMAECS 550

RESULT 3  
 T00355  
 hypothetical protein KIAA0688 - human  
 C:Species: Homo sapiens (man)  
 C>Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 09-Jul-2004  
 C:Accession: T00355  
 R:Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.;  
 DNA Res. 5, 169-176, 1998  
 A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete  
 A:Reference number: Z14142; MUID:98403880; PMID:9734811  
 A:Accession: T00355  
 A:Status: preliminary; translated from GE/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-837 <ISH>  
 A:Cross-references: UNIPROT:O75173; EMBL:AB014588; NID:g3327189; PIDN:BAAJ1663.1; PID:g3:  
 A:Experimental source: brain  
 C:Genetics:  
 A:Gene: KIAA0688  
 F:519-575/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 40.3%; Score 2129; DB 2; Length 837;  
 Best Local Similarity 49.2%; Pred. No. 5.6e-135;  
 Matches 412; Conservative 132; Mismatches 224; Indels 70; Gaps 18;

QY 36 LLLIAALLAVSDALGHPSEDEELVVPKL---ERVPGHGT-TRL---RLHAFDQQLDLV 89  
 DB 37 LLLILLASLLPSARLASPLPREBEIVPPEKINGSLVPGSGTTPARLLCRLOAFGETLLEL 96  
 QY 90 PPSSFLAPGFTLQNVGRK---SGSDTPLPDLAHCFYSGTVNGDPSSAAALSLCEG- 144  
 DB 97 EQDSGVQVGEVTVQJGQAPPELLGGAE---PGT-----YLTGTINGDPESVASLHWGGA 148  
 QY 145 VRGAFYLLGDAYFTQPLPAASERLATAPGKPPAPLQFHLRRNRQDVGGTGCVVDE 204  
 DB 149 LLGVLYRGAEHLQPLEGGTPNSA-GPGCA-----HILRRK----- 184  
 QY 205 PRPTGKAETDEDEBETGEDEGEPQWSPQDPALQGVQPTGTGSIKRRKFRVSSHRYVETML 264  
 DB 185 -----SPASGGQGPCNVKAP-----LGSPSPRPR-RAKRFASLSRFVETLV 224  
 QY 265 VADOSMAEFHSGGLKHVLLTLFSVAALRYKHPISIRNSVSLVWVKILVIHDBOKPEVTSN 324  
 DB 225 VADDKMAAFHAGLURKRYLLTVMAAAAKAFKHPISIRNPVSLVTVTRVLILGSGEBGFQGPS 284  
 QY 325 AALTRNFCNWQKHNPSPDSQDAEHYDTAILFTRQDLICGSQTCDTLGMADVGVTCDFSPRS 384



Db 285 AAQTLRSFCAMGRLNTPSDSDPHFDTAITLFTRODLGCVSTCDTLGMADVGVCDPARS 344  
QY 385 CSVIEDDGLQAAFTTAHELGHVFNPHDDAKOCASING-VNODSHMMASMLNSLDHNSOPW 443  
Db 345 CAIVEDDGLQAAFTTAHELGHVFNHLDHNSKPCISLUNGPLSTSRHMAPVMAHVDPPEPW 404  
QY 444 SPCSGYMTSFLDNGHGECLMDKPNPIQLPGDLFGTSYDANRQCQFTFGEDSKHCPDAA 503  
Db 405 SPCSARFITDFLNGVGHCLLKPAPLHLPTVTFPGDYADARQCQLTFGPDSPHCPQLP 464  
QY 504 STCSILWCTGTSGGVLVVCTQKHPHWDGTSCEGKWCINGKCNVKNRKHFTPFHGSWG 563  
Db 465 PPCAALWCSGHLNHAMCQTKHSPWADGTPCPAPACMGGRCLHMDQLDQDENIPQAGGW 524  
QY 564 MNGPWGDCSRTCGGVQVYTMRECDNPPKNGGKYCEGKRVYRSNLEDCPDNNGKTFRE 623  
Db 525 PWGPGWDCSRTCGGVQVFSRRDCTVPVPRNGKYCEGRTRFRSNTEDCTPGSALTFR 584  
QY 624 ECEAHNFSKASFGSPA-VWIPKYAGVSPKDRCKLICQAKGIGYFVLPQKVVVDGTP 682  
Db 585 EQCAAYNHRDIL-FKSFPGMDVPRYTGVAPODQCKLTCQARALGYVYVLEPRVVDGTP 643  
QY 683 CSPDSTSVQCQCVKAGCDRIIDSKKPKDKCGVCGGNGSTCKISGVSFSAPGYHDI 742  
Db 644 CSPDSSSVQCQRCIHAGCDRIIGSKKPKDKCMWCGDGGSGSGSFRKFRYGNVNV 703  
QY 743 TIPTGATNIEVKORNGSRNNGSFLAIIKAAAGTVILNGDYTLSTLEODIMYKGVV-LRY 801  
Db 704 TIPTGATHILVROQNGPHRS--IYALAKLPDGSVALNGEYTLMPSPDVLPGAVSLRY 761  
QY 802 SGSSAALERSFSPLKEPLTIQVLTGVGNALRPKIYTYFVKKKKESFNAIPTFSAW 859  
Db 762 SGATAASETLGSHGLPLAQPLTIQVLVAGNPQDRLRYSFVPRTPS-TPRPTPDWL 818  
3  
RESULT 4  
T21371  
hypotheical protein F25H8.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequences\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T21371; T24896  
R:Gajadsky, S.  
submitted to the EMBL Data Library, February 1996  
A:Reference number: Z19413  
A:Accession: T21371  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2165 <WIL>  
A:Cross-references: UNIPROT:Q19791; EMBL:Z69360; PIDN:CAA93287.1; GSPDB:GN00022; CESP:F25H8  
A:Experimental source: clone F25H8  
R:Gajadsky, S.  
submitted to the EMBL Data Library, February 1996  
A:Reference number: Z19949  
A:Accession: T24896  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2165 <W12>  
A:Cross-references: EMBL:Z69361; PIDN:CAA93288.1; GSPDB:GN00022; CESP:F25H8.3  
A:Experimental source: clone T13H10  
C:Genetics:  
A:Gene: CESP:F25H8.3  
A:Map position: 4  
A:Introns: 31/1; 52/1; 135/2; 193/3; 216/1; 266/1; 495/2; 547/3; 584/3; 634/2; 744/1; 811/1  
Query Match 30.2%; Score 1598; DB 2; Length 2165;  
Best Local Similarity 32.0%; Pred. No. 8.4e-99;  
Matches 355; Conservative 164; Mismatches 355; Indels 236; Gaps 29;  
QY 48 DALGRPSEE-----DEELVPELERVPVGHGT-TRLR-----LHAFDQQLDLDVDPD 92  
Db 78 DSYGRGRKRKDVTSRRRRRLOQVARGDCGHCLRRLRSDDAVYIVLHRWNQ-----IPDS 132

QY 93 SSFLAPGFTLQNGR-----KSGSDTPIPETDLAHCFYSGTGVNGPSSAAALSL 141  
Db 133 HNKSVPHPSNGNFAPMVLYLDSBEEVRGCMRTPDP-----CIYRAHVKG-VHQHSIVNL 186  
QY 142 C---EGVRGAYLLGEAYFIQPLPAASERLATAAAGEKPPAPLOPHLLRRNRQDVGTC 198  
Db 187 CDSBDGLYGMIALPSGIHTVEPIISGNGTEHDGASRHRQ-----HLVRK----- 230  
QY 199 GVWDDPRPTKABTEDEDEGEDEGPQWSPDPAQLQGVQGTGTSIRKRFVSS-H 257  
Db 231 ----PDPMHFKSFDHLNSTSVNETETIVATWQDQ-----WEDVIERKASRRASNSWD 279  
QY 258 RYVETMLVADOSMAEFHSGUGLKHLLTLFSVAARLYKHPISRNSVSLVVKILVHDBQK 317  
Db 280 HYVEVLVVADTKMYEYHRSLELYTLFSTVASIYRHSRLASINVVVVKLIILKTENA 339  
QY 318 GPEVTSNAALTRNFCWQKOHNPSPDRDAEHYDTAILFTRODLGSGQ-TCDTLGMADV 376  
Db 340 GPRITQNAQTLQDFCRWQYINDPDDSSVQHDVAIILLTRKDCRSQKCDTLGLAELG 399  
QY 377 TVCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACASLNGVQ----- 425  
Db 400 TMDWQKSCAIIEDNGLSAAFTIAHELGHVFSIPHDDERKCSYMPVNVKCKFQSTKFDK 459  
QY 426 -----DSHMMASMLNSLDHSPSPCSGYMTSFLDNGHG--ECLMDKPNQPIQLP 475  
Db 460 TQFNQNFHIMATPLEYNTHPMSWSPCSAGMLERFLENRGTQCLFDQPVERRRYEDVFV 519  
QY 476 -DLGCTSDANRQCQFTFGEDSKHCPDAASTCSLTWCTGTSGGVLVCTQKHPWADGTSC 534  
Db 520 RDEPGKYDAHQCKFVFGPAELCP-YMPTCRLWCATFYGSGQMGCKTQIMPWADGTPC 578  
QY 535 GEGK--WCINGKCNVKNRKHFTDPFHGSGWGMGPWGDSCSRTCGGVQVYTYMRECDNPPK 592  
Db 579 DESRMFCHGACVRLAPESL--TKIDGQWDMRSGECSTCGGVQKGLRDCDSPKPR 636  
QY 593 NGKYCEGKRVYRSNLEDCPDNNGKTFREEQCEAHNEFSKASFG-----SGPAVEWTPK 648  
Db 637 NGKYCVGQRYRSCNTQCEPMDT-QPYREVQC---SEFNKDIGIQGVASTNTHWVVK 692  
QY 649 YAGVSPKDRCKLICQAKGIGYFVLPQKVVVDGTPCSPDSTSVQCQCVKAGCDRIIDSK 708  
Db 693 YANVAPNERCKLYCRLSGSAAPYLLRDKVVDGTPCDRNGDDICVAGACMPAGCDHQLHST 752  
QY 709 KFKDKCGVCGGNGSTCKKISGVSFSAPK-GYHDIITITPTGATNIEVKORNGSRNNGSF 767  
Db 753 LRDKCGVCGGDDSSCKVVKGTFTNEQGTFGYNEVWKIPAGSANIDIRKGYNNMKEDDNY 812  
QY 768 LAIKAAADGTYLNGDYTLSTLEODIMYKGVVLYRSGSSAALERISFSPLEPLTIQVLT 827  
Db 813 LSLRAANGEFLNGHFQVSLARQQTAFQDVTLEYSGSDAIIERINGTGPISRSDIYVHVLS 872  
QY 828 VGNALRPKIYTYF----- 841  
Db 873 VGSH-PPDISYEYMTAAVPNAVIRPISALYLRVTDVTWTECDRACRQOQKMLCLDMS 931  
QY 842 -----VKKKESFNA--IPTFSAMVIBEWGECSSKSCELGMORRIVECRDINGQ 887  
Db 932 THROSHDNCNVLPKPKQATRCMNCIDCSTRITVEDVSSCSAKCGSGQKQRQVSCVMEDG 991  
QY 888 ---PASE----- 891  
Db 992 RQTPASEHLCDRNSKPSDIASICYSGRKWNYGWTSCSETCSGNGKMRKSKYCVDDSN 1051  
QY 892 -----CAKEVKPASTPCADHPCQWQLGEMWSSCKTCGKYKTKSLKCLSHDGVLS 944  
Db 1052 RVDSELGCRGQKATERECNRIPCPRWVYGHWSECSKSGGVMRHAQCLDAADRETH 1111  
QY 945 HDSCDPLKPKXK-----FIDFCTMAECS 967  
Db 1112 TSCGPAQTQEHCHNEHACTWQFGVWSDCS 1141

```
RESULT 5
T18517
procollagen N-endopeptidase (EC 3.4.24.14) I - bovine
N:Collagen names: procollagen N-proteinase
C:Species: Bos primigenius taurus (cattle)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18517
R:Collage, A.: Nussgens, B.V.; Lapiere, C.M.
submitted to the EMBL Data Library, February 1996
A:Description: Cloning of the cDNA of the bovine procollagen I N-proteinase.
A:Reference number: Z18941
A:Accession: T18517
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1205 <COL>
A:Cross-references: UNIPROT:P79331; EMBL:X96389; NID:e990769; PID:e228215; PIDN:CAA65253
A:Experimental source: skin
C:Genetics:
A:Gene: PC I-NP
C:Function:
A:Description: catalyzes cleavage of the propeptides of type I and II collagens prior to
C:Keywords: hydrolase; metalloproteinase

Query Match      20.7%; Score 1096; DB 2; Length 1205;
Best Local Similarity 31.4%; Pred. No. 2.1e-65;
Matches 308; Conservative 144; Mismatches 394; Indels 136; Gaps 44;

QY 6 PGFG-----RRKLGDMGNAERAP-GSRSGFVPVPTLLALLAAALLAVSDALG 51
DB 41 PQGHGAERILAVPVRTDAQGLRVSHVSAATAPAGVTRRAAP-----AQIPGLS 90
QY 52 RSEDEELVPELSEVGHGTRTLRLHAFQOQLDVPDSSFLAPGFTLQNVGRKSGS 111
DB 91 GSEED-----PG-GRLFNVTVFGRDLHLRLPNARLAVAPATVEWQGESAT 138
QY 112 DTPLPETDLAHCFYSGTVNG-DPSSAAALSLCEGVRGAFYLLGEAYFTQPLPAASERLAT 170
DB 139 RV--EPLGLTCLYGVDAEASSVALSNCDGLAGLRMEEEFFIEPL-----EKGLA 191
QY 171 AAPGKPPAPLQFHLRLNRQDVGTCGVVDDEPRPTGKAETEDDEGTGEDEGPQWS 230
DB 192 AKAEARQGRVHVYHPTTSR-----PPPLG-----GPOALDTGISAD 228
QY 231 PQDPALQGVG--OPTGGSIRKRFVSSHRY-VETMLVADQSMARHGS-GLXHYLLTL 285
DB 229 SLDSLSRALGVLEERVNSRRMRRAADDDYNIEVLGVDDSVVDFGTEHVQKYLTL 288
QY 286 FSAARLYKHPSIRNSVLVVKILVI-HDEQKGPEVTSNAALTILNFCNW-QKQHNPPS 343
DB 289 MNIVNEIYHDESIGAHINVLVRIILSVGKMSLIEIGNPSQSLNVCRAWYLQKQPD 348
QY 344 DRDAHYDTALLFTRODL--CGSQTCDTLGMADVTCVDPSSRCSVIEDDGLQAAFTAH 401
DB 349 DHD-EYHDHAIFLTRQDFGPGMQ-----GYAPVTGMCHPVRSCTLHNEHGFSSAFVVAH 402
QY 402 ELGHVFNPHD-DAKQASLNGVNQDSHMAASMLNLDHSPQSPCSGYMLTSLFDNGHG 460
DB 403 ETGHVLGMEHQGNRCG--DEVRLGS--IMAPLVOAARFHWRSRCSQOELSRVLS--Y 457
QY 461 ECLMDKP--QNPQLQFGLPSTYDANRQCOFTFGEDSKHCP--DAASTCTSLMCTGTSG 516
DB 458 DCLRDDPPTHWDPALP-QLPGLHYSMNEQCRDFGLGYMMCTAFRTDPCKQLWCSPDN 516
QY 517 GVLVCOTKHFPWADGTSCEGKWCINGKCVNKNRKHFDTP-----FHGSGWGMGPGWDCS 572
DB 517 PYF-CKTKKGPPLDGTMCAPGKCFKGHCI-----WLTEDILKRDGNWGAWSPFGSCS 568
QY 573 RTCGGGVQYTWRECDNPVPKNGKVCGRVYRSCNLEDPCDNNKTFREEOCEAHNEF 632
DB 569 RTCGTGVKFRTRQCDNPHANGERTCSGLAVDFQLNSQDCPDALA-DFREEQCRQWDLY 627
QY 633 SKASFGSGPAP-EWIPKYAGVSPKDRCKLICOAKGIGYFFVLQPKVVDGTPCS-PDSTSV 690
```

```
Db 628 ----FEHGDAQHWLP-HEHEDAKERCHLYCESKETGEVVSMMKRMVHDGTRCSYKDAFSL 682
QY 691 CVQGCQVAGCDRIIDSKKPKDKGCVCGNGNSTCKKISGVS TSA--KPGYHDIITIPGA 748
Db 683 CVRGDCRVGCDGVIGSSKQBDKCGVCGGDNHSHCVKVGKTFSRSPKLGKLYIKMFEIPAGA 742
QY 749 TNEIVKORNGRSRNNGSFLAIKAAD-GTYILNGDYTLTLEQDQIMYKGVVLYRYSGSA 807
Db 743 RHLIQUAD-----TSHHLAVKMLETGKFLINENDVDPNKSTFIAMGVWEVDEDR 797
QY 808 LERISFSPKLEPLTIQVLTVGNALRPKIKYTFVKKKK-----ESFNAIPTFSA---WVIE 861
Db 798 -ETLQTMGFLHGTITVLVIPEGDA-RISLTVKYMIHEDSLNVDNNDVLEDDSGVGEWALK 855
QY 862 EWGECSSKSCELGWORRLVECR---DINGQPASECAKEVKP-ASTRPCADHPC--PQWQLG 915
Db 856 KWPSCSKPCGGGSOFTYGCRRRLDHRMVRHGFCDVSVKPKAIRTRCTNPQSCSPVWVTG 915
QY 916 EWSSCSKTCGK-GYKTSLSKCL 936
Db 916 EWEPCSRSCGRGTGMQVRSRCV 937

RESULT 6
C89114
protein C37C3.6a [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: C89114
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:1981916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.eleg
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: C89114
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1558 <STO>
A:Cross-references: UNIPROT:Q81710; GB:chr_V; PIDN:AAC25867.1; PID:g3294501; GSPDB:GN000;
C:Genetics:
A:Gene: C37C3.6a
A:Map position: 5

Query Match      12.2%; Score 645; DB 2; Length 1558;
Best Local Similarity 31.3%; Pred. No. 5.5e-35;
Matches 153; Conservative 73; Mismatches 163; Indels 100; Gaps 18;

QY 546 VKNKHKHFDTPPHGSGMGMWPGDCSRTCCGGVQYTMRECDNPVPKNGGKYCEGKRVY 605
Db 68 VDKSGQKE-----TGNWGPWVPENECSRSCGGGVQLEKRCQSGD-----CTGASVRY 114
QY 606 RSNLEDCPDNNKGTFRBEOCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICOAK 665
Db 115 ISCNLNAAC--ESGTDFAEQCSKFNDAL----DGNVHKWTP-YKG---KNKCELVCKE 164
QY 666 GIGYFFVLQPKVVDGTPCSPDSTSVCVQGCVAGCDRIIDSKKPKDKGCVCGNGSTCK 725
Db 165 SGNFYKWAQVVDGTRKCDKSKNDICVDGECLPVGCDGLGSLKFDKCDKCDGDSGTC 224
QY 726 KISG--SVTSAKPGYHDIITPTGATNIEVKQRNQRNNGSFLAIKAADGTYILNGDY 783
Db 225 TIEGRFDERNLSPGYHDIKLPAGATNIKQE--ARKSTNN---LALKNSDHFYFLNGN- 278
QY 784 TLTLEODIMYKGVVLYRYSGSSAALRIRSFPLKEPLTIQVLTVGNALRPKIKYTFVK 843
Db 279 GLIQVEKEVEVGGTIFYV--DDAEPETLSAQGPLSELTVALLPFRKGRSDTAIKYBFSIP 336
QY 844 KKKE-----SF 849
Db 337 LEBEVDYMYKFDNWTPCSVSCGKGVQTRNLYCIDGRNKGVRVEDDLCEENNATKPEFEKSC 396
QY 850 NAIPTFSAWIEBWGECSKSC-ELGWQRRLVECRDI--NGQPAS-----ECAKEVKPASTR 902
```

Db 397 ETVDCAEWFTGDMWSSCSCTGCGQYRVVYCHQVFANGRRVTVEDGNCCTVE-RPPVKQ 455  
 QY 903 PCADHPCPQWQGEWSSCSKTCGKGYKTKSLKLS-----HDGVLSHSDCDPLKPKPHI 958  
 Db 456 TCNRFACPEWQAGPWSACSEKCGDAFQYRSVTCRSEKEGEGKLLAADACPADEQEKFT 515  
 QY 959 D-FCMAEC 966  
 Db 516 ERTCNLGPC 524  
 RESULT 7  
 T34395  
 hypothetical protein C37C3.6b - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T34395, T34394  
 R:Geisels, C.; Bradshaw, H.  
 submitted to the EMBL Data Library, July 1996  
 A:Description: The sequence of C. elegans cosmid C37C3.  
 A:Reference number: Z21518  
 A:Accession: T34395  
 A:Status: preliminary; translated from GB/EMBL/DBBJ  
 A:Molecule type: DNA  
 A:Residues: 1-2167 <GEI>  
 A:Cross-references: UNIPROT:O76840; EMBL:U64857; PIDN:AAC25868.1; GSPDB:GN00023; CESP:C3  
 A:Experimental source: strain Bristol N2; clone C37C3  
 A:Accession: T34394  
 A:Status: preliminary; translated from GB/EMBL/DBBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1555, 'SKP' <GE2>  
 A:Cross-references: EMBL:U64857; PIDN:AAC25867.1; GSPDB:GN00023; CESP:C37C3.6a  
 A:Experimental source: strain Bristol N2; clone C37C3  
 C:Genetics:  
 A:Gene: CESP:C37C3.6b; CESP:C37C3.6a  
 A:Map position: 5  
 A:Introns: 32/3; 104/2; 156/2; 207/3; 459/2; 536/3; 577/2; 1105/3; 1367/1; 1438/1; 1556/

Query Match 12.2%; Score 645; DB 2; Length 2167;  
 Best Local Similarity 31.3%; Pred. No. 8.3e-35;  
 Matches 153; Conservative 73; Mismatches 163; Indels 100; Gaps 18;  
 QY 546 VKNHRKHFDTPFHSGWGMWPGDCSRCTCGGVQVYTMRECDNPVPMKNGKCYCEGRVRY 605  
 Db 68 VDKSGQKE-----TGNWGPWPVENECSRCGGVQLEKRCQSGD-----CTGASVRY 114  
 QY 606 RSCNLEDCPDNGKTFREOCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAK 665  
 Db 115 ISCNLNAE--ESGTDPRABQCSKFNDAL-----DGNHYHKWTP-YKG---KNKCELVCCKPE 164  
 QY 666 GIGYFVLQPKVVDGTFPCSPDSTSVCQGCYKAGCDRIIDSKKFDKCGVCGGNGSTCK 725  
 Db 165 SGNFYKWKADKVDGTGKCDKSKNDICVDGECPLVPGCDGKLGSLKFDKCKGCDGSGSTCK 224  
 QY 726 KISG--SVTSAPKGYHDIIITPGATNIEVKORNGSRNNSFLAIAKAADGTYLNGDY 783  
 Db 225 TIEGRFDERNLSPGYHDIIKLPAGATNIQES--ARKSTNN--LALKNGSDHDFVLNGN- 278  
 QY 784 TLSTLEQDIWYKGVVLRYSGSSAALERISFPLKEPLTIQVLTGVGNALRPKIKTYFVK 843  
 Db 279 GLIQVEKEVEGGTIFVY--DDAEPETLSAQGLSELTVALLPKGRSDRTAIKYEFSIP 336  
 QY 844 KKKE-----  
 Db 337 LEEEDVYMYKFDNWTFCVSGKGVQTRNLICIDGKNKGRVEDDLCEENNATKPFEEKSC 396  
 QY 850 NAIPTFSAWVIEWEGECSKSC-ELGQRRRLVECRDI--NGQPAS-----ECAKEVKRPASTR 902  
 Db 397 ETVDCAEWFTGDMWSSCSCTGCGQYRVVYCHQVFANGRRVTVEDGNCCTVE-RPPVKQ 455  
 QY 903 PCADHPCPQWQGEWSSCSKTCGKGYKTKSLKLS-----HDGVLSHSDCDPLKPKPHI 958

Db 456 TCNRFACPEWQAGPWSACSEKCGDAFQYRSVTCRSEKEGEGKLLAADACPADEQEKFT 515  
 QY 959 D-FCMAEC 966  
 Db 516 ERTCNLGPC 524  
 RESULT 8  
 T18856  
 angiogenesis inhibitor homolog - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T18856; T24653  
 R:McMurray, A.  
 submitted to the EMBL Data Library, July 1995  
 A:Reference number: Z19031  
 A:Accession: T18856  
 A:Status: preliminary; translated from GB/EMBL/DBBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1444 <WIL>  
 A:Cross-references: UNIPROT:Q8MYA8; EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN00028; CESP:C0  
 A:Experimental source: clone C02B4  
 R:McMurray, A.  
 submitted to the EMBL Data Library, July 1995  
 A:Reference number: Z19917  
 A:Accession: T24653  
 A:Status: preliminary; translated from GB/EMBL/DBBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1444 <W12>  
 A:Cross-references: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GN00028; CESP:C02B4.1  
 A:Experimental source: clone T07C5  
 C:Genetics:  
 A:Gene: CESP:C02B4.1  
 A:Map position: X  
 A:Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 566/3  
 Query Match 11.6%; Score 614.5; DB 2; Length 1444;  
 Best Local Similarity 23.8%; Pred. No. 5.6e-33;  
 Matches 237; Conservative 121; Mismatches 346; Indels 291; Gaps 45;  
 QY 62 VPELVRVPG-----HGTRLRHLHAFQOQLDLDVPPDSSFLAPGFTLQNVGRKSGSD--T 113  
 Db 26 VPEVSLIEATRHPLKNGLNKMKFTAWNDTYHLNLRKRSRVSP--HIISVVVRHGDDVTT 83  
 QY 114 PLPETDLAHCFYSYGVNGDPSSAALSCEGVGAFYLLGAYFIQPLPAASERLATAAP 173  
 Db 84 YAGLRDYEQCHYQGEVSKSHGNKAAISDCGALMSIVMEDHFLVLQTLPK----- 133  
 QY 174 GEKPPAPLQFHLLRRNRQDVGTCGVVDDEPRPTGKAETDEDEGEGEPQWSP-- 231  
 Db 134 -----RVHHLQKERRHLVYKRSAGLLTN-----AESKIRBEITRLQEEQESFCDTS 178  
 QY 232 --QDPALQ-----GVGQPTGTG-----SIRKRFVSS-----HRYVETML 264  
 Db 179 EQLDDPWTIPAHLHFNVTIPTSALQDSSFTFPNMDPITLIGLFLDSKLFHEHREVIQ 238  
 QY 265 VADQSMAEFHGSLGHYLLTLFSVAARLYKHPISRNVSLSVVKILVIHDE-----QKQPE 320  
 Db 239 DAEGHLLFEFS-----LALINNHHVLYQQDTLTPNLDIVIVYEMWRTQPSALLSTGVH 290  
 QY 321 VTSNALTLRPNWQKQHNPPSD-RDAEHVDTAILFTFQDLCSGQTCOTGLMADVGTV 379  
 Db 291 KNGQAQSLDLDAFCRYQAHMNFETDLTMNHYDHGVLVLLTGYDIYHT-TTTSVAGVAPVARM 349  
 QY 380 DPSRSCSVIEDGQAAFTTAHELGHVFNMPHDDAK-OC-----ASLNGVNODSHMA 431  
 Db 350 DPLFACSLVEGLHGRSFVLAHEHGNMGMVHDGVQNCNKGCCCLMSAVNGAKTIT----- 405  
 QY 432 SMLSNDLSQWSPSCGYMITSFLL-----DNHGCECLMDKPNQIQLPG-----DL 477  
 Db 406 -----WSDCSVREFNAFLQLQDESGRGNCLRDAS-----PGLISTNHLSDLRL 448  
 QY 478 PGTSDYDANRQOFTFGEDSK-HCPDAA-----TCSTLWCTGTSGGVLCQTKHFPWADGTS 533

Db 449 PGORFTADQCSYFNGRDYKVEIPNGKAMDDICRIILWC-GNSGSTI--STAHPALEGSW 504  
QY 534 CSEGKWCINGKCVNKNRKHFP-----DTPFHSGWGWG-----PWGDCSRCTGGGVQYT 582  
Db 505 CGANKWCHKGQCT-----HWTFGLTVPVIDGESEWGAEGKCPICQCAVSGSITVQOQ 558  
QY 583 MRECDNPVKNKGKCEGKRVRYSNLEDCPDNNGKTFREBQCEAHNEFSKASFGSPA 642  
Db 559 HRDCVNPAPNNGKTCGEGNIRGIVC-----GAT--SSNCLG-----FTRFEEFGN--- 601  
QY 643 VEWIPKYAGVSPKRCCKLIQAAGIGYFFVLQP-----KVVDGTPGSP 685  
Db 602 -----KICSSIKYDPHKPDQQLTGEGFHSHTQPCRWCHLIGSELI RNKGQFPDTPGCF 656  
QY 686 DSTSVCGOQCVKAGCDR--IIDSKKFKDKCGVCGNGSTCKKISGVS TSAKPGVH--DI 741  
Db 657 D--AYCVGQCCLALSCDNKALVEQPE-----DCPRIEGRSVHOWEESWSE 701  
QY 742 ITIPTGATNIEVKORN-ORGSNRNGSFLAIAKAADGTIYLNGDYTLSTLEQDIMYKGVVLR 800  
Db 702 CSVSCGLGREVREKCSGRKCGQVSESRPEG-----VLR 739  
QY 801 YGSSAALERIRSFSPLEPLTIQVLTGNALRPKIYTYFYKKKESFNAIPTFSAWVI 860  
Db 740 -----DCEEFGEW-- 747  
QY 861 BEWGECSKSELGWORLVECRDINGQPASECAKEVKPASTRPCADHPC-POWQLGEWSS 919  
Db 748 KEWGSCEKCALGVQKFRPC--LTDQCSSHLQE-----BRPCDNEGCTNWD--EWSS 798  
QY 920 CSKTGCGKYKTSKLSLASH--DGGVLSHSDCDPLK 952  
Db 799 CSQSCGGRRYIRKCLDDKCDKDDLEXESCNTQK 833  
RESULT 9  
T16892  
hypothetical protein T19D2.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: T16892  
R:Bentley, D.  
submitted to the EMBL Data Library, December 1995  
A:Description: The sequence of C. elegans cosmid T19D2.  
A:Reference number: Z18599  
A:Accession: T16892  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-860 <BEN>  
A:Cross-references: UNIPROT:Q22580; EMBL:U42846; NID:g1125809; PID:g1125810; PIDN:AAA836  
C:Genetics:  
A:Gene: CESP:T19D2.1  
A:Introns: 56/3; 96/1; 171/3; 199/3; 245/2; 293/2; 335/3; 395/3; 426/3; 484/2; 505/3; 58  
Query Match 8.8%; Score 466; DB 2; Length 860;  
Best Local Similarity 23.5%; Pred. No. 2.7e-23;  
Matches 180; Conservative 100; Mismatches 319; Indels 168; Gaps 37;  
QY 248 IRKRFVSSHR-----YVETMLVADQSM-----AEPHSGSLKHYLLTLFSVAAR-----LYKH 295  
Db 119 VKDIRRAEPHRRDILTVELAVFADDAMWDFKMYKAAEENMHTFIMAVVNNIDVLVYTQ 178  
QY 296 PSIRNSVSLVVKVLVI-----HDEQKGPEVTSNAALTIRNFCNWKQKHNPSPSDRDAHYD 351  
Db 179 RLLQPRINIKTVRYEILKNIPLHLMNARKHSGDVRLLDAFCQYQNEINPNPDADPRWD 238  
QY 352 TAILFTRODLGSCOTCDTLGMADVTCVDPKRSKSVLEDDGLQAAFTTAHELGHVFNWPH 411  
Db 239 HALLFSGYDLHRNGVKTVAGYAPVKMCSGVRSCTINEGLDFGVSFVVVTHMGHSLGMYH 298  
QY 412 DDAQKQASLNGWQDSH-----MMASMLSNLDSQPSWPCSGYMITSFLDNGH-----GE 461

Db 299 DGDNEC-----DLRCCIMSPSVSGKTH---WSQCSVNMATFV--GHLGDFFRPPN 345  
QY 462 CLMDKPNQPIQL-----PGDLPGTSYDANRQCQFTFGEDSKHCPPAAST-----CSTLWCTG 513  
Db 346 CLQDASANEQRMVAFKESPEPQLFTLDBQCEBIFHGBCWKHELKDGQTMQNICQWVNC-G 404  
QY 514 TSGGVLCVOTKHPFADGTCGCKWCINGKCVNKNRKHFPDTPFHSGWGWGPGWD--CS 572  
Db 405 NGEVVI--RTAH-PALEGTTCGFGMIRCQGCVCSSQLMRVT-----GGWSTWDRPA 455  
QY 573 RTCGG-----GVQYTMRECDNPVKNKGKCEGKRVRYSNLEDCPDNNGKTF 621  
Db 456 PTCGRCSCQCEIRQIRIMRSIROCNNPSSNNGGAPCQGDGEARGVMVCHRDVNCNGDSIENY 515  
QY 622 REEOCEAHNEFSKA--SFGSGPAVEWIPKYAGVSPKRCCKLIQAAGIGYFFVLQPKVVD 679  
Db 516 ATRVCSRLRDENAIPTILSSEGMOF-----EQAMCKIWCILISGSTNIRTYS-NFPD 566  
QY 680 GTPCSPDSTSVCGOQCVKAGCDRIIDSKKFKDKCGVCGNGST-----CKKISG 729  
Db 567 GAPCGPG--QYCIKGECPILC-----GSTTLAYSEADCPLSVLQT 605  
QY 730 SVTSAKPGYHDIITPTGATNIEVKORNQGRSNGSFLAIAKAADGTIYLNGDYTLSTLE 789  
Db 606 TTTTPPMHVHSVDQF-AGKTN-PYKEHKTPFLNWSGWSVWSRCVTY----- 651  
QY 790 QDIMYKGVVLYRSGSSAAL--ERIRSFSPLEPLT--IQVLTVG--NALRPKI----- 836  
Db 652 -DCHTQGVKVRVRRCLAGVCAGALRERQPCPTGSRPLTTSPPOOTFRNRFIAPLPN 710  
QY 837 KYTVFVKKKESFNAIPTFSAWVIEEGECSKSELGWQ--RRLVECRDINGQPASECAKE 895  
Db 711 RQTNMILRKVDHWG-----PMSACSVTCGTGQKLRRENC---TGO---ECAA- 752  
QY 896 VKPASTRPCADHPCPQ-----W-QLGEWSSCSCTGKGYKTSKLSKLS 937  
Db 753 -----TGPCVMQSCRENKNTWTEWSQWSDCVNCGEGVQFRKACFA 794  
RESULT 10  
T22545  
hypothetical protein F53B6.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T22545  
R:White, S.  
submitted to the EMBL Data Library, October 1996  
A:Reference number: Z19578  
A:Accession: T22545  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-1059 <WIL>  
A:Cross-references: UNIPROT:P90884; EMBL:Z81086; PIDN:CA803121.1; GSPDB:GN00019; CESP:F5;  
A:Experimental source: clone F53B6  
C:Genetics:  
A:Gene: CESP:F53B6.2  
A:Map position: 1  
A:Introns: 38/3; 92/3; 131/3; 169/1; 236/2; 282/3; 349/2; 376/2; 420/1; 576/3; 759/3; 84;  
Query Match 7.9%; Score 418.5; DB 2; Length 1059;  
Best Local Similarity 24.6%; Pred. No. 5.5e-20;  
Matches 102; Conservative 66; Mismatches 121; Indels 125; Gaps 16;  
QY 561 SWGMWGPWGDSCSTCGGVQYTWRECDNPVKNKGKCEGKRVRYSNLEDCPDNNGKT 620  
Db 25 SWAAWSPWSSCTTKCGGVSRQLRCLT-----SKCSGESVRFKVC----- 65  
QY 621 FREEOCEAHNEFSKASFGSGPAVEWIPKYAGVSPKRCCKLIQAAGIGYFFVLQPKVVDG 680  
Db 66 -AQKTESKSLARDITCGEEI-----VS-RGOCEVVCRSRLTGANFLW--RVDDG 113  
QY 681 TPC-SPDSTSVCGOQCVKAGCDRIIDSKKFKDKCGVCGNGSTCKKISGVS TSAKPGYH 739

Db 114 TPCQATSRVCSKSGCQIVGCDGLISSRFRDFAACVCGGRGDTCT----- 158

Qy 740 DIITPTGATNIEVKORNGRRNNNGSFLAIKAADGTGYILNGDYTLSTLEQDIWYKGVL 799  
||| : : : : :  
: ||| : : : : :  
159 -----DNCKFI-----WKVEEYTACASNC-----DIV 182

Qy 800 RYSGSAALERIRSFPLKEPLTIQVTV-----GNALRPKIYTFVKKKESF 849  
||| : : : : :  
: ||| : : : : :  
Db 183 DWSGAG-----RSIASTSQPIVVVCNAITGRVVPEKLCADKLRPKVE-----A 225

Qy 850 NAIPTF--SAWVIBEWGECSKCELGWQRRLVECRDINGO-----PASECAKEVKBPAST 901  
||| : : : : :  
: ||| : : : : :  
Db 226 RPCPMLICSRMAWAUETECVPHCGEGTRREVYCQTAAHNVTVHPDTFCENGTRPAAE 285

Qy 902 RPCADHPCCQWLGESSSKTCGKYGKTSLKCLSHDGGVLSHDSCDPLKKPK 955  
||| : : : : :  
: ||| : : : : :  
Db 286 ENCUSTSCGRWEAGKWSKCTASCQGVRHRHVACV--GG----SDCDEGGRPR 332

RESULT 11  
Tt5976  
hypothetical protein F08C6.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: Tt5976  
R;Bentley, D.  
submitted to the EMBL Data Library, June 1995  
A:Description: The sequence of C. elegans cosmid F08C6.  
A:Reference number: Z18440  
A:Accession: Tt5976  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-957 <BEN>  
A:Cross-references: UNIPROT:Q19204; EMBL:U29378; NID:G668184; PID:G668185; PIDN:AAA68721  
X:Experimental source: strain Bristol N2  
C:Genetics:  
A:Gene: CESP:F08C6.1  
A:Introns: 23/1; 135/3; 220/3; 285/3; 325/2; 376/3; 461/2; 498/3; 532/3; 627/2; 714/1; 714/1;

Query Match 7.6%; Score 402; DB 2; Length 957;  
Best Local Similarity 21.3%; Pred. No. 6.2e-19;  
Matches 227; Conservative 122; Mismatches 384; Indels 354; Gaps 53;

Qy 78 LHAFQQLDLDPD--SSFLAPGFTLNVGKSGSDTPLPET-DLAHCYFGTVNGDPSS 135  
:||| : : : : :  
: ||| : : : : :  
Db 61 IQAFNKYNLSLEXTLAKLLSSGVTVVKNKGGSLDFGSTLDSCHVHYG-----EKV 115

Qy 136 AAALSCEG-----VRGAFYLGEAYFTQPLP---AASERLATAAPEKPPAPLQ 182  
||| : : : : :  
: ||| : : : : :  
Db 116 YAAISNCGRIVSESNRKQKGTVIDDGEIIVVHPFDHHAHRKRAPEGA----- 167

Qy 183 FHLL--RRNRQDVGCCTGVVDDEPRPTGKAETED---EDGETGEDSGPQSPDPAQ 237  
||| : : : : :  
: ||| : : : : :  
Db 168 -HVYKRETLAGEKDFCGUL-----DNVTVESLVEDESAIPEDV----- 206

Qy 238 GVGPQTGTGSIRKKRFVSRR-----YVETMLVDAQSM-----AEFHGS---GLKHYL 282  
||| : : : : :  
: ||| : : : : :  
Db 207 -----FVTGORLTQQSLDI VELA VFVDENLWRHFSSKGGGMADRKLQDYT 251

Qy 283 LTLFSVAARLYKHPSTRNSVSVLVVVKILVIHDEOKGPEVTS-----NAALTURNFCNW 335  
||| : : : : :  
: ||| : : : : :  
Db 252 LTLNNIQIMTYQPTASPPLTFVRIVEVL---TRQPSALAGYLHNHGNAQMFLDRFCRY 308

Qy 336 QKQHNPSPDRDAEHYDTAILFTQDL---CGSQTCDDLTMADVCTCDPSSCSVIEDDG 392  
||| : : : : :  
: ||| : : : : :  
Db 309 QRN---LAVRDWDH---AIMLTGYDIRGAGSRIS--GIARLDGMCDFMTCTLAEGLD 360

Qy 393 LQAATTAHELGHVFNMHPDDAK-----QCASLNGVNO-----DS 427  
||| : : : : :  
: ||| : : : : :  
Db 361 FTSAFIGTHELGHRTTELKDHSIDLFTGTGCGPSKWCOLGRCVPWTGNEIOPTVQVHA 420

Qy 428 HMMASMLSNLDSQPWS-----PCSGYMITSFLDNGHBELMDKPQNPIQLPGDL 477  
||| : : : : :  
: ||| : : : : :

421 PVVTLPRIIDGS--WSWGATICSQCTCNGILGSLGLAIARTCSAPYPANG---GSDC 475

478 PGTSYDA---NRQOFTFEEDSK-----HCPDAASTCTLTWCTGTSGVLVQCQKHF 526

476 VGSTRAVLCRQC---GRASKSVDEYISDKMEQKRLKNDRELTKGS-----QLNRF 526

527 P-----WADTSCGEGKWCINGCKCVNKH-- 553

527 PQRACKVPDVOQHYGSQRNYRFFGDNLPDGTSCGYDRYCLDGECLALNCNNALISRQ 586

554 -----FDTPEFGSGMGMWPGWDCSRTCGGGOVYTMRECDNPVPGKNGKCYCEGK 601

587 SCTPDTCTPITQSSSVYRGQWTLWLTCTATCGGGRKRNACISIT-----GQCEGN 640

602 RVYRSCNLEDCPDNNNGKTFREEQCEAHNEFSGKSGGPAVEWIPKYAGVSPKDRCKLI 661

641 EDETEVCSSSCP-----SVLRVGNWSTWT-----EW-----NHCSVS 674

662 C-----QAKGIGYFFVLQPKVDGTCSPDS-----TSVCVOGQCVKACCDR 703

675 CGRGSQAR---YRKCLSPHRTLAFCPGENKVTNELRIITFFKARSYIMCSVRCKIKRNT 731

704 IIDSKKPKDC-----GVCG--GNGSTCKKISGVSATKPGYHDIITPTGATNIEV 753

732 ISEKNIEVRSDNGPCNAIGWGTGGWSTCS-----TSCGPG-----TL-----V 772

754 KQRNQRGRNNGSF-----LAIKAAADGYILNGDYTLSTLEQDIMYKGVLRVYSGSSA 806

773 RQRTCNREPCDGAHERSCNVATCQNDGIWSLWNEWSDCS---RVCGKGL----- 820

807 ALERIRSPSLKEPLTIQVLTVGNNALRPKIKTYTFVKKKKESNAIPTSAWIEWGECEC 866

821 ---RSRSRS-----CFGSGCMGASSPQFCNEQACASSANDWGTW---SGWSQC 864

867 SKSCELGHQRRLVECRDINGQPASCAKEVKASTPACADHPCPO-----WQLGEMSS 919

865 SVSCAGVRRKTRTCTGTG-----CPNGYKESAI--CNDRCENKNAWGWG--GYWSS 914

920 CSKTCGKGKKTSLKCLSHDGGVLSDHSCDPLKKPKHFDICTMAEC 966

915 CSETCGDGVRRVRKC-----YSGNCGDGOQYK---QYCNLRVC 951

RESULT 12

T00260

hypothetical protein KIAA0605 - human

C/Species: Homo sapiens (man)

C/Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 09-Jul-2004

C/Accession: T00260

R/Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O. DNA Res. 5, 31-39, 1998

A/Title: Prediction of the coding sequences of unidentified human genes. IX. The complete

A/Reference number: Z14086; MUID:96290545; PMID:9628581

A/Accession: T00260

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-951 <NAG>

A/Cross-references: UNIPROT:O60345; EMBL:AB011177; NID:G3043733; PIDN:BAA25531.1; PID:G303

A/Experimental source: brain

C/Genetics:

A/Note: KIAA0605

F/46-106/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 7.4%; Score 391; DB 2; Length 951;

Best Local Similarity 22.8%; Pred. No. 3.4e-18;

Matches 134; Conservative 61; Mismatches 194; Indels 200; Gaps 21;

Qy 562 WGMWGPWDCSRTCGGQVYTMREC-----DNPVPGKNGKCYCEGKRVYRSCNLEDCPDNN 617

Db 50 WGETKWTAFSRSCGGVTSQERHCLQQRKSVPGPNRTCTGTSKRYQLCRVQCECP-D 108

Qy 618 KGTFREEQCEAHNEFSGKSGGPAVEWIP-----KYAGVSPKDRCKLIQAKGIGYFFVL 673

Db 109 GRSFEEQCVSFN-----SHVYNGRTHQWKPLYPDDYVHISSKP-CDLHCTTVD-QQRQLM 162  
Qy 674 QPKVVDGTFPCS-PDSTSVCVQOCVKAGCDRIIDSKKFKDFKCGVCGNGSTCKKISGSVT 732  
Db 163 VP-ARDGYSCKLTDLRGVCVSKCFPIGCDGVLFSTHLDKCGICQGDGSSCTHVTGNRY 221  
Qy 733 --SAKPGYHDIITPTGATNIEVKORNGSRNNGSFLAIAAADCTYILNGDYTLSTLEQ 790  
Db 222 KGNALHGLSVLTHIPAGARDIOIVER-----KKSADVLALADEAGYFFNGNYKVDSPK 275  
Qy 791 DIMYKGVVLRVSGS-----SAALERIRSFSPKLEPTIOVLTVGNALRPKIKYTPVKKKK 846  
Db 276 NFNIAGTVVKYRRPMDVYETGIEYVAOQPTNOGLNMVMVN-QNGKSPSITFEYTLLOPP 334  
Qy 847 ESFNAIPFSAWVIEE-----WGECSSCELGHQRLV----- 879  
Db 335 HESRPQIYYGFSESAESQGLDAGLGMFIPHNGSLYQOAS-SERLGLDNRLFGHPGLDM 393  
Qy 880 -----EC-----RDIN----- 885  
Db 394 ELGPSQOETNEVCEQAGGACEGPPRGKGRDRNVGTPLTGDKDDEVDTHFASQEPF 453  
Qy 886 -----GQPASECAK----- 894  
Db 454 SANAIISDQLLAGSLDKDFTLNETVNSIFAQAPRSSLAESFFVDYBENEGAGPYLLNGS 513  
Qy 895 -----EVRPASTRPCAD-----HPCP--QWOLGEW 917  
Db 514 YLELSSDRVANSSEAPPFNVTSLTSGNRTHKARTRPKARKOGVSPADMYRWKLJSSH 573  
Qy 918 SSCSTCKGKYKTSKLCSLSHDGSDDLKPKKHFTDFCTMAEC 966  
Db 574 EPCSATCTGWSAYAMCVRGDVEVDSDYCDALTRPEPVHEFCAGREC 622  
RESULT 13  
S60257  
meltin alpha - mouse  
A:Species: Mus musculus (house mouse)  
C:Date: 10-Apr-1996 #sequence\_revision 19-Apr-1996 #text\_change 09-Jul-2004  
R:Accession: S60257  
R:Yagami-Hiromasa, T.; Sato, T.; Kurisaki, T.; Kamiyo, K.; Nabeshima, Y.I.; Fujisawa-Seh Nature 377, 652-656, 1995  
A:Title: A metalloprotease-disintegrin participating in myoblast fusion.  
A:Reference number: S60257; MUID:96026308; PMID:7566181  
A:Accession: S60257  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-903 <YAG>  
C:Cross-references: UNIPROT:O61824; EMBL:D50411; NID:G1054586; PIDN:BAA08912.1; PID:G105421-503/Domain: disintegrin homology  
C:Superfamily: mouse meltin alpha; disintegrin homology  
F:421-503/Domain: disintegrin homology <Dis>  
F:349/Active site: Glu #status predicted  
Query Match 7.2%; Score 378.5; DB 2; Length 903;  
Best Local Similarity 22.0%; Pred. No. 2.2e-17;  
Matches 189; Conservative 102; Mismatches 278; Indels 289; Gaps 37;  
Qy 21 AERAPGRSFGVPVPTLLLAALIAVSDALG-----RPSDEELVPELER 67  
Db 6 ARRAPPARAL-----LLALAGALLAPRAARGMSLWDQRGAYEVARASILLSDKDPGTPG-QS 59  
Qy 68 VPGHG-----TTRLRLHAFDQDLDDVPDSSFAPGFT-----LQNVGRKSGSDTLPET 118  
Db 60 IPAKDHPDLVTQQLSERD--LILSLERNGLIANGFTETHYLDQ-----GTDVSLTRN 112  
Qy 119 DLACHFYSGTVNGDPSSAAALSLCEGVGRGAFYLGEAYFIQPLPAASERLATAAPGEKPP 178  
Db 113 HTDHCYHYGHVQGDAAASVVSLSTCSDLRLGFENKTYSLRPMKNTD----- 160  
Qy 179 APLOPHLLRRNRQGDVGTCGVVDDEPRPTGKAETDEDEGEDEGPQWSPQDPALQG 238  
Db 161 ---SYKLVPAESMTNIOQLCG-----SQHNKSLNTMEDVSP----- 193

Qy 239 VGQPTGTGSIIR-----KKRFVSSSHRYVETMLVADQSMAEFHGSG-----LKHYLLTLFSPA 289  
Db 194 -----GTSQMARRHKRETLKMTKYVELVIVADN--REFQRGKDLEKVKQORLIEIANHV 246  
Qy 290 ARLYKHPISIRNSVLVVKILVHDEQKGPVETVNAALTLRNFNCNQWQHPPSDRDAEH 349  
Db 247 DKFYRPLNIR-----IVLVGVEVWMDIK-CSISQDPPFTRLHFEFLDWRKIKLLP-----RKS 297  
Qy 350 YDTAILFTRQDLCSQTCDTLGMADVGVTCVDPSPSCSVI-----EDDGLQAAKTTAHELGHV 406  
Db 298 HDNAQLISGVYFOGT-----TIGMAPIMSMCTAEOGGVMDHSDSPGLAAVATLAHELGN 353  
Qy 407 FNMPHDDAKOCASLNGVNDSHMMASMLNLDHSQSPSCGYMITSPFLDNHGHCLEMDK 466  
Db 354 FGMNHDILRCSCSRMAAEKGCIMNFSTGPPFMVFFSSCSRKLEASLEKMGMCLENL 413  
Qy 467 PQNPIQLPGLDLPGTSDYDANRQCQFTFGEDSKHCPDAASTCSTLWCTGTGTGGVLVCQTKHF 526  
Db 414 PEVQAQFG-----RKCGNGYVEEGEC----- 436  
Qy 527 PWADGTCGEGKWCINGKCVKN-----HRKHFDTPFHGSGWGMGPGWDCSRTC 575  
Db 437 -----DCGEPEECTN-RCCNATTCITLKPAVCAH-----GQCCEDC 471  
Qy 576 -----GGGVQVYTMRECDNPVPKNGKYCEGRVRYRSCNLEDCPDN----- 616  
Db 472 QLKPPGTACRGSSNSCDLP-----EFTCG-----TAPHC PANVYLHDGHPCOQVDG 517  
Qy 617 ---NGKTFREBQCBNAHFEKASFGSGPAVEWIPKYAGVSPKCKLICQAKGIYGFVL 673  
Db 518 YCYNG-----ICQTHEQOCQVTLWPGAK-----PAPGICFERVN 551  
Qy 674 QPKVVDGTPCSPDSTSVCVQOCVKAGCDRIIDSKKFKDFKCGVCGNGSTCKKISGSVTS 733  
Db 552 SA-----GDP-----YGNCK-----DSKSAFAKCEL---RDAKCGKIQCOGGA 587  
Qy 734 AKPGYHDIITPTGATNIEVKORNGSRNNGSFLAIAAADGTYYLNGDYTLSTLEQDIM 793  
Db 588 SRP-----VIGTNAVSIETNIPQEGGR-----ILCRGTHVYLG-----DMP 625  
Qy 794 YKGVVLRVSGSSAALERI 811  
Db 626 DPGLVL--AGTKCAEGKI 641  
RESULT 14  
S48169  
metalloproteinase (EC 3.4.24.-) H-II precursor - carpet viper (fragment)  
N:Contains: disintegrin  
C:Species: Echis pyramidum leakeyi  
C:Date: 06-Dec-1996 #sequence\_revision 06-Dec-1996 #text\_change 09-Jul-2004  
R:Accession: S48169  
R:Paine, M.J.I.; Moura-Da-Silva, A.M.; Theakston, R.D.G.; Crampton, J.M. Eur. J. Biochem. 224, 483-488, 1994  
A:Title: Cloning of metalloproteinase genes in the carpet viper (Echis pyramidum leakeyi).  
A:Reference number: S48169; MUID:95010025; PMID:7925363  
A:Accession: S48169  
A:Molecule type: mRNA  
A:Residues: 1-549 <PAI>  
C:Cross-references: UNIPROT:Q90500; EMBL:X78971; NID:G763094; PIDN:CAA55566.1; PID:G763094  
C:Superfamily: mouse meltin alpha; disintegrin homology  
C:Keywords: hydrolase; metalloproteinase; venom; zinc  
F:1-138/Domain: propeptide (fragment) #status predicted <PRO>  
F:139-549/Product: metalloproteinase H-II #status predicted <MAT>  
F:350-432/Domain: disintegrin homology <Dis>  
F:284/Active site: Glu #status predicted  
Query Match 6.6%; Score 347; DB 2; Length 549;  
Best Local Similarity 23.0%; Pred. No. 1.5e-15;  
Matches 152; Conservative 80; Mismatches 206; Indels 222; Gaps 32;  
Qy 106 GRKSGSDTLPETDLAHCFYSVGTVNGDPSSAAALSLCEGVGRGAFYLGEAYFIQPLPAAS 165

Db 38 GREITNPVED-----HCYHGRVQNDASHSSASISACNGLKGFKLQGGTYFIEPLKI-- 91  
Qy 166 ERLATAAPGKPPAPLQPHLLRRNRQDVGCGVVDDEPRPTGKAETEDDEDEGTGEDE 225  
Db 92 -----PDSEAHAYKVENTEKEDQ 110  
Qy 226 GPQ-----WSPQDPALQGVQPTGTSIRKRFVSSHRYVETMLVADQSMAPFHG-- 275  
Db 111 APQMGCVTHTNMESDEPIKEA---SRLVASSEQQSYDNFRVYKLVIVVDHMRVTKYND 167  
Qy 276 SGLKHYLLTLFSAARLYKHPISRSVSLVVKI-----LVHDEQKGPVETSNAALTIR 330  
Db 168 STIRTRYEMVNTVNEIYTHAIR--VALVLEFWSNGDLI-----NVTSAAEHTLN 217  
Qy 331 NFCNMOKQNPSPD-RDAEHYDTAILFTQDLGSGTCTDLGMADVGTVCDSRSCSVIE 389  
Db 218 LFGVWR-----ASDLLSRHRDHAHLTAIDLNG-----PTIGLRDVSSWCQATRSVGVQ 268  
Qy 390 DDG---LQAAFTTAHELGHVFNMPHDDAKQCASLNGVNDQSHMMASMLSNLDHSPWSPC 446  
Db 269 DHSPTVRAVAVTMAHEMGNLGNH-DGNHC---NCGANSICMAAVLRN-PAPEYFSDC 322  
Qy 447 SGYMTSFLDNGHGBCLMDKPNQPIQLPGDLGTSVDANRQCFTEGDSKSKCPDAASTC 506  
Db 323 SRRYQNFNLTYIPDCTIIRPK-----TDIVSPQVCGNGLLDGEBC-DCGSPA 371  
Qy 507 STLWCTGTSGGVLVCOTKHPFADGTSCGEGKW--CINGKCVNKNHRKHFDPFHGSGM 564  
Db 372 N-----CO---YPCDAASCKLHWSVECFHCCDQCFKPACTECRG----- 411  
Qy 565 WGPWGDSCSRTCGGVQVYTMRECDNPVPKNGKGYCEG-----KRVRYRSCNLE 611  
Db 412 -----IRSECDLP-----EYCTGQSAECLRMFSTRWTKMPKLSLLIQ 448  
Qy 612 -DCPDNNGKTFREQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRC-KLICQAKGIOY 669  
Db 449 WTCP-----IMGYOCYAH-----FGQNAVY-----GQDACFEINKEGKGDPY 485  
Qy 670 -----PFVLQPK-----VVD-GTPCSPDSTSVCVQGCQV 697  
Db 486 CRKENDVPIPCAQEDVKCGRLCFETEPNMCRIYVPGDEGWDVGTGKC--EDKKVCINGKCI 543

RESULT 15  
S24789  
jararagin C precursor - jararaca (fragment)  
N:Alternate names: single chain botrocetin  
N:Contains: disintegrin-like 28k protein; hemorrhagic proteinase (EC 3.4.24.-)  
C:Species: Bothrops jararaca (jararaca)  
C>Date: 20-Feb-1995 #sequence revision 29-Aug-1997 #text\_change 09-Jul-2004  
C:Accession: S24789; J2245; A4463; A37958; J22373  
R:Paine, M.J.I.  
submitted to the EMBL Data Library, August 1992  
A:Reference number: S24789  
A:Accession: S24789  
A:Molecule type: mRNA  
A:Residues: 1-571 <PAI>  
A:Cross-references: UNIPROT:P30431; EMBL:X68251; NID:G62467; PID:G62468  
R:Usami, Y.; Fujimura, Y.; Miura, S.; Shima, H.; Yoshida, E.; Yoshioka, A.; Hirano, K.;  
Biochem. Biophys. Res. Commun. 201, 331-339, 1994  
A:Title: A 28 kDa-protein with disintegrin-like structure (jararagin-C) purified from B  
A:Reference number: J2245; MUID:94256999; PMID:8198592  
A:Accession: J2245  
A:Molecule type: protein  
A:Residues: 360-571 <USA>  
A:Experimental source: venom  
R:Paine, M.J.; Desmond, H.P.; Theakston, R.D.; Crampton, J.M.  
J. Biol. Chem. 267, 22869-22876, 1992  
A:Title: Purification, cloning, and molecular characterization of a high molecular weight  
ily.  
A:Reference number: A4463; MUID:93054601; PMID:1385408  
A:Accession: A4463

A:Molecule type: mRNA  
A:Residues: 1-23, 'Q', '25-92, 'G', '94-131, 'G', '133-169, 'Q', '171-571 <PA2>  
A:Cross-references: GB:X68251  
A:Experimental source: venom gland  
A:Note: sequence inconsistent with the nucleotide translation  
R:Fujimura, Y.; Tlicani, K.; Usami, Y.; Suzuki, M.; Oyama, R.; Matsui, T.; Fukui, H.; Sug  
Biochemistry 30, 1957-1964, 1991  
A:Title: Isolation and chemical characterization of two structurally and functionally di  
A:Reference number: A37958; MUID:91129280; PMID:1993206  
A:Accession: A37958  
A:Molecule type: protein  
A:Residues: 360-372, 'E', '374-378, 'X', '380-384 <FUI>  
A:Note: 361-Val was also found  
C:Comment: Inhibits collagen- and ADP-induced platelet aggregation.  
C:Superfamily: mouse meltrin alpha; disintegrin homology  
C:Keywords: hydrolase; metalloproteinase; venom; zinc  
F:360-571/Product: jararagin C #status experimental <MAT>  
F:362-444/Domain: disintegrin homology <DIS>  
F:295,299,305/Binding site: zinc (His) #status predicted  
F:296/Active site: Glu #status predicted

Query Match 6.5%; Score 343; DB 2; Length 571;  
Best Local Similarity 21.9%; Pred. No. 3e-15;  
Matches 158; Conservative 94; Mismatches 228; Indels 242; Gaps 36;

Qy 48 DALGRPSDEBELVPELERVPGHGTTRLRHLAFDQDLDDVPPDSSFLAPGFTLQNVGR 107  
Db 14 DAMQVEFKVNGEPPVHLKKNKG-----LFSKQYS-EIHVSPD-----GR 52  
Qy 108 KSGSTPTLPETDLAHCFSYGVNGDPSSAALSCEGVARGAFYLLGEBAYFTPLPAASER 167  
Db 53 EITTYPPVED---HCYYHGRINDADSTASACNGLKGYFKLQRETYFIEPLK---- 104  
Qy 168 LATAAPGKPPAPLQPHLLRRNRQDVGCGVVDDEPRPTGKAETEDDEGTGEDEGP 227  
Db 105 -----PDSEAHAYKVENVEDAP 125  
Qy 228 Q-----WSPQDPALQGVQPTGTSIRKRFVSSHRYVETMLVADQSMAPFHGSLKH 280  
Db 126 KMGVGTQNWKSVEP-IKKASQLAFTA--EQQRY-DPKYIBFFVVVDGQTVTKNNGDLD- 180  
Qy 281 YLLTLFSAARLYKHPISRSVSLVVKILVIHDEQKGPVTSN-----AALTIRN 331  
Db 181 -----KIKARMYELANIVNE---IFRYLYMHVALVLEIWSNGDKITVPRDVEDYTLNS 230  
Qy 332 FCNMOKQNPSPD-RDAEHYDTAILFTQDLGSGTCTDLGMADVGTVCDSRSCSVIED 390  
Db 231 FAEWRK-----TDLLTRKKNDAQLLTAIDFNG-----PTIGYAYIGSMCHPKRSYGIQD 281  
Qy 391 ---DGLQAAFTTAHELGHVFNMPHDDAK-QCASLNGVNDQSHMMASMLSNLDHSPWSPC 446  
Db 282 YSPINLVAVIMAHMGNLGIHHDGTGSCSGDYPCI-----MGPTISN-EPSKFFSNC 334  
Qy 447 SGYMTSFLDNGHGBCLMDKPNQPIQLPGDLPGTS-YDANRQCFTEGDSKH-CPDAAS 504  
Db 335 SYIQCWDFIMNHNPECIINEPLGTDIISPPVCGNELLEVGEBCDCTPENQNECCDAAT 394  
Qy 505 TCSTLWCTGTSGGVLVCOTKHPFADGTSCGEGKWCKINGKCVNKNHRKHFDPFHGSGM 564  
Db 395 -----CKLK-----SGSQCH----- 405  
Qy 565 WGPWGDSCSRTC-----GGGVQVYTMRECDNPVPKNGKGYCEGKRVYRSC-----NLED 612  
Db 406 ---GDCCEQCKFSKSGTECRASMSECD-----PAEHCTGS-----SECPADVPHKNGQP 452  
Qy 613 CPDNNKGTFR-----EQCEAHNEFSKASFGSGPAVEWIPKYAG 651  
Db 453 CLDNYGYCYNGCPTMYHQCYALFGADVYEADSCFDKNQ-----KGYNYGYCKRENG 505  
Qy 652 ---VSPKD-RC-KLICQAKGIG-----YFFVLQPK-----VVDGTPCSPDSTSVCVQGO 695  
Db 506 KKIPCAPEDVKGRUYCKDNPSFGQNPNCKMFYSNDDHKGWVLPGTKCA--DGKVCNSNGH 563

Qy 696 CV 697  
||  
Db 564 CV 565

Search completed: August 22, 2005, 10:19:38  
Job time : 53 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2005, 10:06:04 ; Search time 179 Seconds  
(without alignments)  
2766.372 Million cell updates/sec

Title: US-09-989-687-126  
Perfect score: 5287  
Sequence: 1 MORAVPEGFRKLGSDMGN.....CDPLKPKHFIDFCTMAECS 967

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_03.\*  
1: uniprot\_9prot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5232	99.0	967	1	AT51_HUMAN	Q9ubi8 homo sapien
2	5234	98.8	967	2	Q8NE26	Q8ne26 homo sapien
3	4293.5	81.2	968	1	AT51_MOUSE	P97857 mus musculus
4	4293	81.2	967	2	Q68EJ2	Q68ej2 rattus norv
5	4287	81.1	967	1	AT51_RAT	Q9wug1 rattus norv
6	3902.5	73.8	759	2	Q8RZM8	Q8hzm8 equus cabal
7	2477.5	46.9	950	1	AT15_HUMAN	Q8te58 homo sapien
8	2320.5	43.9	890	1	AT58_HUMAN	Q9up79 homo sapien
9	2265.5	42.9	905	1	AT58_MOUSE	P57110 mus musculus
10	2124	40.2	837	1	AT54_HUMAN	O75173 homo sapien
11	2117	40.0	837	2	Q6UW8	Q6uwa8 homo sapien
12	2106	39.8	837	2	Q6F4Q8	Q6p4q8 homo sapien
13	2105	39.8	833	2	Q8K384	Q8k384 mus musculus
14	2104.5	39.8	845	2	Q8BNJ2	Q8bnj2 mus musculus
15	2075.5	39.3	839	2	Q7YS95	Q7ys95 bos taurus
16	1979	37.4	630	1	AT54_RAT	Q9esp7 rattus norv
17	1976	37.4	893	2	Q6A017	Q6a017 mus musculus
18	1924	36.4	930	1	AT55_HUMAN	Q9una0 homo sapien
19	1911.5	36.2	930	1	AT55_MOUSE	Q9r001 mus musculus
20	1904.5	36.0	928	2	Q67Y19	Q6ty19 rattus norv
21	1862	35.2	1306	1	AT20_MOUSE	P59511 mus musculus
22	1825	34.5	1935	1	AT59_HUMAN	Q9p2n4 homo sapien
23	1794.5	33.9	867	2	Q6SKM3	Q66km3 xenopus tro
24	1727.5	32.7	1311	1	AT20_HUMAN	P59510 homo sapien
25	1672	31.6	623	2	Q8BGP4	Q8bgp4 m mus muscu
26	1614	30.5	562	1	AT15_MOUSE	P59384 mus musculus
27	1598	30.2	2165	2	Q19791	Q19791 caenorhabdi
28	1359	25.7	1641	2	Q68SA9	Q68sa9 mus musculus
29	1321	25.0	1686	2	Q6P7J9	Q6p7j9 homo sapien
30	1280	24.2	1593	1	AT12_HUMAN	P58397 homo sapien
31	1279.5	24.2	1221	2	Q6P4R5	Q6p4r5 homo sapien

RESULT 1									
AT51_HUMAN									
ID	AT51_HUMAN	STANDARD;	PRT;	967	AA.				
AC	Q9UH18; Q9NSJ8; Q9P2K0; Q9UH83; Q9UP80;								
DT	30-MAY-2000 (Rel. 39, Created)								
DT	16-OCT-2001 (Rel. 40, Last sequence update)								
DE	ADAMTS-1 precursor (SC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1) (METH-1).								
DE	Name=ADAMTS1; Synonyms=KIAA1346, METH1;								
OS	Homo sapiens (Human)								
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
OX	NCBI_TaxID=9606;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RA	Casas C., Pritchard M.A., Estivill X., Arbones M.L.;								
RT	"Cloning, characterization and mapping on human chromosome 21 of the orthologue of murine Adamts-1.;"								
RL	Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.								
RN	[2]								
RP	SEQUENCE FROM N.A., AND FUNCTION.								
RC	TISSUE=Heart;								
RX	MEDLINE=99367466; PubMed=10438512; DOI=10.1074/jbc.274.33.23349;								
RA	Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S., Lombardo M., Iruela-Arispe M.L.;								
RT	"METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new family of proteins with angio-inhibitory activity.;"								
RL	J. Biol. Chem. 274:23349-23357(1999).								
RN	[3]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Endothelial cells;								
RX	MEDLINE=20247184; PubMed=10785405;								
RA	Glienke J., Schmitt A.O., Pilarsky C., Hinzmann B., Weiss B., Rosenthal A., Thierauch K.H.;								
RT	"Differential gene expression by endothelial cells in distinct angiogenic states.;"								
RL	Eur. J. Biochem. 267:2820-2830(2000).								
RN	[4]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Brain;								
RX	MEDLINE=20181126; PubMed=10718198;								
RA	Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;								
RT	"Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro.;"								
RL	DNA Res. 7:65-73(2000).								
RN	[5]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=20289799; PubMed=10830953; DOI=10.1038/35012518;								
RA	Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,								

Q8te57 homo sapien  
Q69z28 mus musculus  
Q8CG28 mus musculus  
Q9h324 homo sapien  
Q8te60 homo sapien  
Q8bky1 m mus muscu  
Q8llb3 mus musculus  
Q9w493 drosophila  
Q7ksh7 drosophila  
Q8sxb0 drosophila  
Q9ukp4 homo sapien  
Q7pw77 anopheles g  
Q8bka1 mus musculus  
Q8k206 mus musculus

RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,  
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,  
RA Shintani A., Sasaki T., Nagamine K., Mitsuama S., Antonarakis S.E.,  
RA Minoshima S., Shimizu N., Nordstieck G., Hornischer K., Brandt P.,  
RA Schaefer M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,  
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,  
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,  
RA Lehrach H., Reinhardt R., Yaspo M.-L.;  
RT Nature 405:311-319(2000).  
RL [6]  
RN SEQUENCE OF 418-967 FROM N.A.  
RC TISSUE=Melanoma;  
RA Blum H., Bauersachs S., Mewes H.-W., Weil B., Wiemann S.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be  
CC involved in its turnover (By similarity). Has angiogenic inhibitor  
CC activity. Active metalloprotease, which may be associated with  
CC various inflammatory processes as well as development of cancer  
CC cachexia. May play a critical role in follicular rupture.  
CC -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1938-Glu-|-Leu-1939  
CC site, within the chondroitin sulfate attachment domain.  
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
CC matrix (By similarity).  
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important  
CC for a tight interaction with the extracellular matrix.  
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By  
CC similarity).  
CC -!- SIMILARITY: Belongs to the peptidase M12B family.  
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.  
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
CC -----  
CC EMBL; AF170084; AAF15317.1; -;  
CC EMBL; AF060152; AAD48080.1; ALT\_INIT.  
CC EMBL; AF207664; AAF23772.1; -;  
CC EMBL; AB037767; BAA32584.1; ALT\_INIT.  
CC EMBL; AP001697; BAA95502.1; -;  
CC EMBL; AL162080; CAB82413.1; -;  
CC PIR; T47158; T47158.  
CC HSP; P07996; 1LSL.  
CC MEROPS; M12.222; -;  
CC Genew; HGNC:217; ADAMTS1.  
CC H-InvDB; HIX0016042; -;  
CC MIM; 605174; -;  
CC GO; GO:0008237; F:metallopeptidase activity; TAS.  
CC GO; GO:0007229; P:integrin-mediated signaling pathway; TAS.  
CC GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
CC InterPro; IPR006586; ADAM\_cysteine.  
CC InterPro; IPR010294; ADAM\_spacer1.  
CC InterPro; IPR001762; Disintegrin.  
CC InterPro; IPR001590; Peptidase M12B.  
CC InterPro; IPR002870; Peptidase M12B.  
CC InterPro; IPR006025; Pept\_M\_Zn\_BS.  
CC InterPro; IPR000884; TSP1.  
CC InterPro; IPR000805; TSP1.  
CC Pfam; PF05986; ADAM\_spacer1; 1.  
CC Pfam; PF01562; Pep\_M12B\_propep; 1.  
CC Pfam; PF01421; Reprolysin; 1.  
CC Pfam; PF00090; TSP1; 3.  
CC PRINTS; PR01705; TSP1REPEAT.  
CC SMART; SM00608; ACR; 1.  
CC SMART; SM00209; TSP1; 3.  
CC PROSITE; PS0215; ADAM\_MEPPO; 1.  
CC PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.

DR PROSITE; PS0214; DISINTEGRIN\_2; FALSE\_NEG.  
DR PROSITE; PS0092; TSP1; 3.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
KW Extracellular matrix; Glycoprotein; Heparin-binding; Hydrolase;  
KW Metalloprotease; Repeat; Signal; Zinc; Zymogen.  
FT SIGNAL 1 49 Potential.  
FT PROPEP 50 252 By similarity.  
FT CHAIN 253 967 ADAMTS-1.  
FT DOMAIN 253 475 Metalloprotease.  
FT DOMAIN 476 559 Disintegrin-like.  
FT DOMAIN 559 614 TSP type-1 1.  
FT DOMAIN 617 724 Cys-rich.  
FT DOMAIN 725 849 Spacer.  
FT DOMAIN 854 905 TSP type-1 2.  
FT DOMAIN 908 967 TSP type-1 3.  
FT DOMAIN 843 846 Poly-Lys.  
FT SITE 198 198 Cysteine switch (Potential).  
FT METAL 401 401 Zinc (catalytic) (By similarity).  
FT ACT SITE 402 402 By similarity.  
FT METAL 405 405 Zinc (catalytic) (By similarity).  
FT METAL 411 411 Zinc (catalytic) (By similarity).  
FT CARBOHYD 547 547 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 720 720 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 764 764 N-linked (GlcNAc...) (Potential).  
FT CONFLICT 227 227 P -> A (in Ref. 4 and 5).  
FT CONFLICT 468 468 Q -> H (in Ref. 1).  
FT CONFLICT 561 561 S -> N (in Ref. 1).  
SQ SEQUENCE 967 AA; 105383 MW; C18938932474LED1 CRC64;  
Query Match 99.0%; Score 5232; DB 1; Length 967;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 957; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 MQRVPSGFGRRKLGSDMGNAERAPGSRSGPVPVTLTLLAAALLAVSDALGRPSEDEEL 60  
Db 1 MQRVPSGFGRRKLGSDMGNAERAPGSRSGPVPVTLTLLAAALLAVSDALGRPSEDEEL 60  
Qy 61 VPELEVRVGHGTRLRHLHAFDQDLDPDSSFLAPGFTLQNVGRKSGDPLPPTDL 120  
Db 61 VPELEVRVGHGTRLRHLHAFDQDLDPDSSFLAPGFTLQNVGRKSGDPLPPTDL 120  
Qy 121 AHCFYSCTVNGDPSSAAALSILCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAP 180  
Db 121 AHCFYSCTVNGDPSSAAALSILCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAP 180  
Qy 181 LQFHLLRRNRQGVGTCGVVDDEPRPTGKAETDEDEGEDEGEDEGEDEGEDEGEDEGEDE 240  
Db 181 LQFHLLRRNRQGVGTCGVVDDEPRPTGKAETDEDEGEDEGEDEGEDEGEDEGEDEGEDE 240  
Qy 241 QPTGTGSIKRRKRVSSHRYVETMLVADQSMAEFHGSLGHVLLTLFVSVAARLYKHPISRN 300  
Db 241 QPTGTGSIKRRKRVSSHRYVETMLVADQSMAEFHGSLGHVLLTLFVSVAARLYKHPISRN 300  
Qy 301 SVSLVVKVILVIHDEQKGPVTSNAALTNRNFCNQKQHNPPSDRDAEHYDTAILFTQD 360  
Db 301 SVSLVVKVILVIHDEQKGPVTSNAALTNRNFCNQKQHNPPSDRDAEHYDTAILFTQD 360  
Qy 361 LCQSQTCDTLGMADVGTVCPSRSCSVIEDDGLQAAFTTAHELGHVFNPMHDDAKQKASL 420  
Db 361 LCQSQTCDTLGMADVGTVCPSRSCSVIEDDGLQAAFTTAHELGHVFNPMHDDAKQKASL 420  
Qy 421 NGVNQDSHMMASMLSNLDHSPQSPSCSGYMITSLDNGHGECMLDKPQNPTQLPGDLPGT 480  
Db 421 NGVNQDSHMMASMLSNLDHSPQSPSCSGYMITSLDNGHGECMLDKPQNPTQLPGDLPGT 480  
Qy 481 SYDANRQCQFTFGEDSKHCPDAAASTCTLMCTGTSGGLVLCQTKHFPWADGTCGEGKWC 540  
Db 481 SYDANRQCQFTFGEDSKHCPDAAASTCTLMCTGTSGGLVLCQTKHFPWADGTCGEGKWC 540  
Qy 541 INKCVNKNHRKHFDTPFHGSGWGWGPGWDCSRTCGGQVQYTWRECDNPPKNGKGYCEG 600  
Db 541 INKCVNKNHRKHFDTPFHGSGWGWGPGWDCSRTCGGQVQYTWRECDNPPKNGKGYCEG 600



QY 661 ICQAKGIGYFFVLPQKWDGTPCSDTSVCVQGCVRAGCDRIIDSKKKFKDCVCGGN 720  
|||||  
Db 661 ICQAKGIGYFFVLPQKWDGTPCSDTSVCVQGCVRAGCDRIIDSKKKFKDCVCGGN 720  
|||||  
QY 721 GSTCKKISGVSPTSAGPKYHDIITPTGATNIEVKQNRQSGRNNGSFLAIKAADGTIYN 780  
|||||  
Db 721 GSTCKKISGVSPTSAGPKYHDIITPTGATNIEVKQNRQSGRNNGSFLAIKAADGTIYN 780  
|||||  
QY 781 GDYTLSTLEODIMYGVNLRVSGSSAALIRISFSPLEKPLTIQVLTGVNLRPKIKYTY 840  
|||||  
Db 781 GDYTLSTLEODIMYGVNLRVSGSSAALIRISFSPLEKPLTIQVLTGVNLRPKIKYTY 840  
|||||  
QY 841 FVKKKESFNALPTFSANWIEBWSCKSCSLGQWRRLVECRDINGQPAASECAKEVKPAS 900  
|||||  
Db 841 FVKKKESFNALPTFSANWIEBWSCKSCSLGQWRRLVECRDINGQPAASECAKEVKPAS 900  
|||||  
QY 901 TRPCADHPCPQWLGEWSCKSCGKGYKTKSLKCLSHDGGVLSHSDCDPLKKPKHFIDF 960  
|||||  
Db 901 TRPCADHPCPQWLGEWSCKSCGKGYKTKSLKCLSHDGGVLSHSDCDPLKKPKHFIDF 960  
|||||  
QY 961 CTMAECS 967  
|||||  
Db 961 CTMAECS 967  
|||||

RESULT 3  
ATSL\_MOUSE  
ID ATSL\_MOUSE STANDARD; PRT; 968 AA.  
AC P97857; O54768;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE ADAMTS-1 precursor [EC 3.4.24.-] (A disintegrin and metalloproteinase  
with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).  
GN Name=Adamts1;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SVJ;  
RX MEDLINE=98110583; PubMed=9441751; DOI=10.1006/geno.1997.5064;  
RA Kuno K., Lizasa H., Ohno S., Matsushima K.;  
RT "The exon/intron organization and chromosomal mapping of the mouse  
ADAMTS-1 gene encoding an ADAM family protein with TSP motifs.";  
RL Genomics 46:466-471(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97150761; PubMed=8995297; DOI=10.1074/jbc.272.1.556;  
RA Kuno K., Kanada N., Nakashima E., Fujiki F., Ichimura F.,  
RA Matsushima K.;  
RT "Molecular cloning of a gene encoding a new type of metalloproteinase-  
disintegrin family protein with thrombospondin motifs as an  
inflammation associated gene.";  
RL J. Biol. Chem. 272:556-562(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP CHARACTERIZATION, AND MUTAGENESIS OF GLU-403.  
RX MEDLINE=99303657; PubMed=10373500; DOI=10.1074/jbc.274.26.18821;  
RA Kuno K., Terashima Y., Matsushima K.;  
RT "ADAMTS-1 is an active metalloproteinase associated with the  
extracellular matrix.";  
RL J. Biol. Chem. 274:18821-18826(1999).  
RN [5]  
RP FUNCTION.  
RX MEDLINE=20389568; PubMed=10930576; DOI=10.1016/S0014-5793(00)01854-8;  
RA Kuno K., Okada Y., Kawashima H., Nakamura H., Miyasaka M., Ohno H.,  
RA Matsushima K.;  
RT "ADAMTS-1 cleaves a cartilage proteoglycan, aggrecan.";  
RL FEBS Lett. 478:241-245(2000).  
RN [6]  
RP FUNCTION, AND INDUCTION.  
RX MEDLINE=20243757; PubMed=10781075; DOI=10.1073/pnas.080073497;  
RA Robker R.L., Russell D.L., Espey L.L., Lydon J.P., O'Malley B.W.,  
RA Richards J.S.;  
RT "Progesterone-regulated genes in the ovulation process: ADAMTS-1 and  
cathepsin L proteases.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:4689-4694(2000).  
CC -!- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be  
involved in its turnover. Has angiogenic inhibitor activity (By  
similarity). Active metalloproteinase, which may be associated with  
various inflammatory processes as well as development of cancer  
cachexia. May play a critical role in follicular rupture (By  
similarity).  
CC -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1691-Glu-|-Leu-1692  
site, within the chondroitin sulfate attachment domain.  
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
matrix.  
CC -!- INDUCTION: Induced in vitro in colon adenocarcinoma cells by  
interleukin-1, or in vivo in kidney and heart by  
lipopolysaccharide. Also induced by LH stimulation in granulosa  
cells of preovulatory follicles.  
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important  
for a tight interaction with the extracellular matrix.  
CC -!- PTM: The precursor is cleaved by a furin endopeptidase.  
CC -!- SIMILARITY: Belongs to the peptidase M12B family.  
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.  
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.  
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a  
frameshift in position 7.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AB001735; BAA24501.1; ALT INIT.  
DR EMBL; D67076; BAA11088.1; ALT\_FRAME.  
DR EMBL; BC040382; AAH40382.1; -.  
DR EMBL; BC050834; AAH50834.1; -.  
DR HSPSP; P07996; ILSL.  
DR MEROPS; M12.222; -.  
DR MGD; MGI.109249; Adamts1.  
DR InterPro; IPR010294; ADAM\_spacer1.  
DR InterPro; IPR001762; Disintegrin.  
DR InterPro; IPR006025; Pept M. Zn BS.  
DR InterPro; IPR001590; Peptidase\_M12B.

```

DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000885; TSP 1.
DR Pfam; PF05986; ADAM spacer1; 1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF00090; TSP1; 3.
DR PRINTS; PR01705; TSP1REPEAT.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.
DR PROSITE; PS0214; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE; PS00992; TSP1; 3.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR Extracellular matrix; Glycoprotein; Heparin-binding; Hydrolase;
KW Metalloprotease; Repeat; Signal; Zinc; Zymogen.
FT SIGNAL 1 48
FT PROPEP 49 253
FT CHAIN 254 968
FT DOMAIN 254 476
FT DOMAIN 477 559
FT DOMAIN 560 615
FT DOMAIN 618 725
FT DOMAIN 726 850
FT DOMAIN 855 911
FT DOMAIN 912 968
FT DOMAIN 195 199
FT SITE 206 206
FT METAL 402 402
FT ACT_SITE 403 403
FT METAL 406 406
FT METAL 412 412
FT CARBOHYD 548 548
FT CARBOHYD 721 721
FT CARBOHYD 765 765
FT CARBOHYD 783 783
FT CARBOHYD 946 946
FT MUTAGEN 403 403
FT CONFLICT 335 335
FT CONFLICT 425 425
SQ SEQUENCE 968 AA; 105841 MW; 42EBDA55499FB6C1 CRC64;

Query Match 81.2%; Score 4293.5; DB 1; Length 968;
Best Local Similarity 81.0%; Pred. No. 5.5e-270;
Matches 790; Conservative 58; Mismatches 112; Indels 15; Gaps 5;

Qy 1 MORAVPEGFGRKLGSDMGNAERAPGRSFGVPVTLTLLAAA---LLAVSDALGRPSEED 57
Db 1 MOPKVPLGRKQKPCSDMGDVORAAARSGLSAHMLLLLLLSITWLLCARGHGRPTSED 60

Qy 58 BELVVPLELRRVPGHG--TTRLRLHAFDQQLDLDVPPDSFSLAPGFTLQNVGRKSGSDTPL 115
Db 61 BELVLPLELRRVPGHDSITTRLRLDAFGQQLHLKLPDQDSGLAPGFTLQTVGRSPGSEAQH 120

Qy 116 --PETDLAHCFYSGTVNGDPSSAAALSICEGVRGAFYLLGEAYFTQPLP-AAERLATAA 172
Db 121 LDPTGDLAHCFYSGTVNGDPGSAALSICEGVRGAFYLLQGEFFITQAPAGVATERLAPAV 180

Qy 173 PGEKPPAPLQPHLLRRNRQDVGCTGVVDDEPRPTGKAETDEDEGTEGEDEGPQWSPQ 232
Db 181 PEESSARPPQPHILRRRRRGSGAKCGVMDDETLP-----SDSRPESQNRQWNPVR 233

Qy 233 DPALQGVGOPTGTGIRKKRFVSSHRYVETMLVADQSMAPFEGSLGKHYLLTFLFVAARL 292
Db 234 DPTPDACKSPGSGIRKKRFVSSPRYVETMLVADQSMADFHGSLGKHYLLTFLFVAARF 293

Qy 293 YKHPISIRSVLVVKILVHDEQGPVTSNAALTLRNFCNQKHPPSPDRDAEHYDT 352
Db 294 YKHPISIRSVLVVKILVHDEQGPVTSNAALTLRNFCNQKHPPSPDRDPDEHYDT 353

Qy 353 ALLFTQDLCCGSOTCDTLGMDADVGVCPDRSCSVIEDDGLQAAFTTAHELGHVFNMPHD 412
Db 354 ALLFTQDLCCGSOTCDTLGMDADVGVCPDRSCSVIEDDGLQNAFTTAHELGHVFNMPHD 413

```

## RESULT 4

```

Q68EJ2 ID Q68EJ2 PRELIMINARY; PRT; 967 AA.
AC Q68EJ2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE A disintegrin and metalloproteinase with thrombospondin motifs 1.
GN Name=Adams1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McWeney P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

```

RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RA Director MSC Project;  
RL Submitted (AUG-2004) to the EMBL/GenBank/DBSJ databases.  
DR EMBL; BC080233; AAH80237.1; -;  
DR GO; GO:0007229; P: integrin-mediated signaling pathway; IEA.  
DR InterPro; IPR006586; ADAM cysteine1.  
DR InterPro; IPR010294; ADAM spacer1.  
DR InterPro; IPR001590; Peptidase M12B.  
DR InterPro; IPR002870; Peptidase M12B\_N.  
DR InterPro; IPR006025; Pept\_N\_Zn\_BS.  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR008085; TSP 1.  
DR Pfam; PF05986; ADAM spacer1; 1.  
DR Pfam; PF01562; Pep M12B propep; 1.  
DR Pfam; PF01421; Repolysin; 1.  
DR Pfam; PF00090; TSP\_1; 3.  
DR PRINTS; PR01705; TSP1REPEAT.  
DR SMART; SM00608; ACR; 1.  
DR SMART; SM00209; TSP1; 3.  
DR PROSITE; PS0215; ADAM\_MPRO; 1.  
DR PROSITE; PS00092; TSP1\_3.  
DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
KW Integrin.  
SQ SEQUENCE 967 AA; 105647 MW; C6349B5D8CBFEA24 CRC64;

Query Match 81.2%; Score 4293; DB 2; Length 967;  
- Best Local Similarity 81.2%; Pred. No. 6e-270;  
- Matches 791; Conservative 59; Mismatches 110; Indels 14; Gaps 5;

Qy 1 MORAVPEGFGRKLGSDMGNAERAPGSRSGFVPTLLLLAAA---LLAVSDALGRPSSED 57  
Db 1 MOPEVPLSGGLKPKCSMDGDIQRAAKFRSSQSAHMLLLLASITWLLCVRGAHGRPTIED 60  
Qy 58 BELVVPPELRLVPGHG-TTRRLRHAPDQQLDLDVPPDSSFLAPGFTLQNVGRKSGSDTPL- 115  
Db 61 BELVPLSLERARGHSTLLRLDADFQQLHLKLPDPSGFLAPGFTLQTVGRSPGSEAOHL 120  
Qy 116 -PETDLACHFCYSGTVNGDPSSAAALSLCEGVRGAFYLLGAEVFIQPLPA-ASERLATAAP 173  
Db 121 DPTGDLACHFCYSGTVNGDPSSAAALSLCEGVRGAFYLLQGEFFIQLPAPAVATERLVPAEP 180  
Qy 174 GEKPPAPLQFHLLRRNRQDYGVTGCGVVDDBPREPTGKAETDEDEGTEGEDEGPQSPQD 233  
Db 181 KEESTAPPRPHILRRRRSGGKGVMDDETLP-----SNSGRESQNTPDQWPLRN 233  
Qy 234 PALQGVQPTGTSIRKRVFVSHRYVETMLVAQOSMAEFHGSGLKHVLLTLFVVAARLY 293  
Db 234 PTPQAGKPTGSGIRKRVFVSSPRYVETMLVADQSWADPHGSGLKHVLLTLFVVAARFY 293  
Qy 294 KHPSTRNSVSLVVKILVIHDEQKPEVTSNAALTNRNFCNWKQHNPPSPDRDAEHYDTA 353  
Db 294 KHPSTRNSISLVVVKILVIHDEQKPEVTSNAALTNRNFCNWKQHNPPSPDRDEHYDTA 353  
Qy 354 ILFTRQDLCSGQTCDTLGMADVGTGCDPFSRSCSVIEDDGLQAFTTAHGLGHVFNMPHDD 413  
Db 354 ILFTRQDLCSGHTCDTLGMADVGTGCDPFSRSCSVIEDDGLQAFTTAHGLGHVFNMPHDD 413  
Qy 414 AKQASLNGVNDSHMAMSLNLDHSQWPSPGSGYMTISPLDNGHCECLMDKQNPQIQL 473  
Db 414 AKHCASFNGVSGDHLMASSLNLDHSQWPSPGSCAYMVTSTPLDNGHCECLMDKQNPQIKL 473  
Qy 474 PGDLPGTSTYDANROCOQTFEGSDSKHCPDAASTCTLWCTGTSGGVLCVCKTHRPWADGTS 533  
Db 474 PSDLPGLTYDANROCOQTFEGSESTHCPDAASTCTLWCTGTSGLLVCCQTKHFPWADGTS 533  
Qy 534 CGEGKWCINGKVCVKNKHKHFDTPPHGSGWGMWGPWGDCSRTCCTGGGVQYTWRECDNPVFN 593

Db 534 CGEGKWCVSGKVCVKNKTDKHFATPVHGSWGPWGPWGDCSRTCCTGGGVQYTWRECDNPVFN 593  
Qy 594 GGGYCEGKRVRYRSCNLEDPCDNNNGKTFREBQCAHNEFFSKASFGSGPAVWIPKYAGVS 653  
Db 594 GGGYCEGKRVRYRSCNLEDPCDNNNGKTFREBQCAHNEFFSKASFGNEFTVETWTPKYAGVS 653  
Qy 654 PKDRCKLICQAKGIGYFFVLQPKVVDGTPCSPDSTSVCSQCVKAGCDRIIDSKKRPDK 713  
Db 654 PKDRCKLTCEAKGIGYFFVLQPKVVDGTPCSPDSTSVCSQCVKAGCDRIIDSKKRPDK 713  
Qy 714 CGVCGNGSTCKKISGVSPTSAPGYPHDIITPTGATNIEVKQRNQRNNGSFLAKAA 773  
Db 714 CGVCGNGSTCKKISGVTSTRPGYHDIPTAGATNIEVKHNRNPRNNGSFLAIRAA 773  
Qy 774 DGTVILNGDYTLSTLEQDIMYKGVVLYRSGSSAALERIRSFSPKPELTTIQTVLTVGNALR 833  
Db 774 DGTVILNGNFTLSTLEQDLYKGTVLYRSGSSAALERIRSFSPKPELTTIQTVLTVGNALR 833  
Qy 834 PKIYTYFVKKKESFNAIPTFSAWVIEEWGECSSKCELGWQRRLVECRDINGOPASECA 893  
Db 834 PKIYTYFVKKKTEPFNAIPTFSEWVIEEWGECSSKTCGSGWQRVVECRDINGHPASECA 893  
Qy 894 KEVKPASTRPCADHPCPQWQLGWSSCSKTCGKYKTKSLKCLSHDGGVLSHSDCDPLKK 953  
Db 894 KEVKPASTRPCADLPQPRWQVGDWSPCSKTCGKYKTKLCLSHDGGVLSHSDCDPLKK 953  
Qy 954 PKHFIDECTMAECS 967  
Db 954 PKHYIDFCILTQCS 967

RESULT 5  
ATSL\_RAT ID ATSL\_RAT STANDARD; PRT; 967 AA.  
AC Q9WUQ1; Q9ER11;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
DE with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).  
GN Name=Adamts1;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
RA Liu X., Tu Y., Yin T., Johnstone E.M., Stephenson D.T., Clemens J.A.,  
RA Little S.P.;  
RT "Induction of a disintegrin and metalloproteinase with the  
RT thrombospondin type I motif (ADAMTS).";  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 18-967 FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Liver;  
RA MEDLINE=20304099; PubMed=10847486;  
RA Diamantis I., Luethi M., Hoessli M., Reichen J.;  
RT "Cloning of the rat ADAMTS-1 gene and its down regulation in  
RT endothelial cells in cirrhotic rats.";  
RL Liver 20:165-172 (2000).  
CC -I- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be  
CC involved in its turnover. Has angiogenic inhibitor activity (By  
CC similarity). Active metalloproteinase, which may be associated with  
CC various inflammatory processes as well as development of cancer  
CC cachexia. May play a critical role in follicular rupture (By  
CC similarity).  
CC -I- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1683-Glu-|-Leu-1684  
CC site, within the chondroitin sulfate attachment domain.  
CC -I- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
CC -I- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
CC matrix (By similarity).

CC -!- INDUCTION: Down-regulated in endothelial cells derived from  
 CC cirrhotic liver.  
 CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important  
 CC for a tight interaction with the extracellular matrix.  
 CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By  
 CC similarity).  
 CC -!- SIMILARITY: Belongs to the peptidase M12B family.  
 CC -!- SIMILARITY: Contains 1 disintegrin-like domain.  
 CC -!- SIMILARITY: Contains 3 TSP type-1 domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF149118; AAD34012.1; -.  
 CC EMBL; AF304446; AAG29823.1; -.  
 CC HSSP; P07996; 1LSL.  
 CC MEROPS; M12.222; -.  
 CC InterPro; IPR006586; ADAM\_cysteine.  
 CC InterPro; IPR010294; ADAM\_spacer1.  
 CC InterPro; IPR001762; Disintegrin.  
 CC InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 CC InterPro; IPR001590; Peptidase\_M12B.  
 CC InterPro; IPR002870; Peptidase\_M12B\_N.  
 CC InterPro; IPR000884; TSP1.  
 CC InterPro; IPR008085; TSP\_1.  
 CC Pfam; PF05986; ADAM\_spacer1; 1.  
 CC Pfam; PF01562; Pep\_M12B\_propep; 1.  
 CC Pfam; PF01421; Reprolysein; 1.  
 CC Pfam; PF00900; TSP\_1; 3.  
 CC PRINTS; PR01705; TSP1REPEAT.  
 CC SMART; SM00608; ACR; 1.  
 CC SMART; SM00209; TSP1; 3.  
 CC PROSITE; PS50215; ADAM\_MEPRO; 1.  
 CC PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 CC PROSITE; PS50214; DISINTEGRIN\_2; FALSE\_NEG.  
 CC PROSITE; PS50092; TSP1; 3.  
 CC PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 CC Extracellular matrix; Glycoprotein; Heparin-binding; Hydrolase;  
 KW Metalloprotease; Repeat; Signal; Zinc; Zymogen.  
 FT SIGNAL 1 54  
 FT PROPEP 55 252  
 FT CHAIN 233 967  
 FT DOMAIN 253 475  
 FT DOMAIN 476 558  
 FT DOMAIN 559 614  
 FT DOMAIN 616 724  
 FT DOMAIN 725 857  
 FT DOMAIN 854 910  
 FT DOMAIN 911 967  
 FT DOMAIN 194 198  
 FT SITE 205 205  
 FT METAL 401 401  
 FT ACT\_SITE 402 402  
 FT METAL 405 405  
 FT METAL 411 411  
 FT CARBOHYD 547 547  
 FT CARBOHYD 720 720  
 FT CARBOHYD 764 764  
 FT CARBOHYD 792 792  
 FT CARBOHYD 945 945  
 FT CONFLICT 21 21  
 FT CONFLICT 26 31  
 FT CONFLICT 49 49  
 FT CONFLICT 72 72  
 FT CONFLICT 79 79  
 FT CONFLICT 249 249  
 FT CONFLICT 262 265  
 FT CONFLICT 607 607

FT CONFLICT 936 936 L -> V (in Ref. 2).  
 FT CONFLICT 962 962 I -> T (in Ref. 2).  
 SQ SEQUENCE 967 AA; 105705 MW; F93C864F6DCDB4CF CRC64;  
 Query Match 81.1%; Score 4287; DB 1; Length 967;  
 Best Local Similarity 81.1%; Pred. No. 1.5e-269;  
 Matches 790; Conservative 59; Mismatches 111; Indels 14; Gaps 5;  
 QY 1 MRAVPEGFGRKLGSDMGAERAPGSRSGFVPVPTLLLLAAA---LLAVSDALGRPSSED 57  
 DB 1 MQPEVPLSGKLPKPCSDMGDIQRAAKFSSQSAHMLLLLLASITMLLCVRGHGRTEED 60  
 QY 58 BELVVPPELERYVGH-GTTRRLRHAPDQOLDLDVPPDSSFLAPFTLQNVGRKSGSDTFL- 115  
 DB 61 BELVLPSELERARGHDSITLLRLDAFGQQLHLKLPDPSGLFAPFTLTQVGRSPGSAOHL 120  
 QY 116 -PETDLAHCFSYGTVNGDPSSAAALSICEGVGRGAFYLLGEAFYIQLPLA-ASERLATAAP 173  
 DB 121 DPTGDLAHCFSYGTVNGDPSSAAALSICEGVGRGAFYLLGEAFYIQLPLA-ASERLATAAP 180  
 QY 174 GEKPPAPLQFHLLRRNRQDVGTCGVVDDEPRPTGKAETEDDEDEGEDEGEPQWSPQD 233  
 DB 181 KEESTAPPRFHLLRRRRREGSGGAKGVNDEETLPT-----SNSGRESQNTPDQWPLRN 233  
 QY 234 PALQGVQPTGTGTSIRKQRFVSSHRYVETMLVAQOSMAEFHSGSLKHYLLTLFVSAARLY 293  
 DB 234 PTPQAGAKPTGPGSIRKQRFVSSPRYVETMLVAQSMADFHSGSLKHYLLTLFVSAARFY 293  
 QY 294 KHPSIRNSVSLVVKILVIHDEOKPEVTSNAALTNRNFCNWKOHNPSPDRDARHYDTA 353  
 DB 294 KHPSIRNSISLVVKILVIYEQKPEVTSNAALTNRNFCNWKOHNPSPDRDARHYDTA 353  
 QY 354 ILFTQDLCSGSQTCDTLGMADVTVCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDD 413  
 DB 354 ILFTQDLCSGSHTCDTLGMADVTVCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDD 413  
 QY 414 AKQASLNGVNDQSHMMASMLSNLDHSPWSPGSGYMITSLFDNGHSGCLMDKPNPQL 473  
 DB 414 AKHCASFNGVSGDHLMASMLSLDHSQWSPGSGYMITSLFDNGHSGCLMDKPNPQL 473  
 QY 474 PGDLPGTSYDANROCOPTFGEDSKHCPDAASTCTSLWCTGTSGGVLCOTKHPWADGTS 533  
 DB 474 PSDLPGLTYDANROCOPTFGBESTHCPDAASTCTSLWCTGTSGGVLCOTKHPWADGTS 533  
 QY 534 CGEGKWCINGKCVNKNHKKHFDTPPHGSGMWGMGDCSRCTCGGVQVYTMRECDNPVKN 593  
 DB 534 CGEGKWCYSGKCVNKTDMKHFPATPHGSGWGPWGDGDCSRCTCGGVQVYTMRECDNPVKN 593  
 QY 594 GGYCEGKRVYRSCNLEDPCDNNNGKTPREOCEAHNEFSSKASFGSGPAVEWIPKYAGVS 653  
 DB 594 GGYCEGKRVYRSCNIEDPCDNNNGKTPREOCEAHNEFSSKASFGNEPTVEWTPKYAGVS 653  
 QY 654 PKDRCKLIQAKGIGYFVLOPKVVDGTPCSPDSTSVCGQCVKAGCDRIIDSKKFKDK 713  
 DB 654 PKDRCKLTCEAKGIGYFVLOPKVVDGTPCSPDSTSVCGQCVKAGCDRIIDSKKFKDK 713  
 QY 714 CGVCGENGSTCKKISGVSVTSAPKGYHDIITPTGATNIEVKORNGSRNNGSFLAIKAA 773  
 DB 714 CGVCGENGSTCKKISGVTSTRPGYHDIITIPAGTNIIEVKHNRPRGSRNNGSFLAIRAA 773  
 QY 774 DGTYLINGDYTLSTLEQDQIMYKVVLYRYSGSSAALERIRSFSPKLEPTITQVLTGVNALR 833  
 DB 774 DGTYLINGNFTLSTLEQDLTYKGTVLYRYSGSSAALERIRSFSPKLEPTITQVLMVGHALR 833  
 QY 834 PKIKYTYFKKKSKSFNAIPTFSAWVIEEWECSEKSCELGHQRLVECRDINGOPASECA 893  
 DB 834 PKIKYTYFMKKKTFFPNAIPTFSEWVIEEWECSEKSKTCSGWQRRVVECRDINGHAPSECA 893  
 QY 894 KEVKPASTRPCADHPCPOWQLGEWSSCSKTCGKYKKTSLKCLSHDGVLSHSDCDPLKK 953  
 DB 894 KEVKPASTRPCADLPFCPRKQVGDWSPCSKTCGKYKKTSLKCLSHDGVLSHSDCDPLKK 953  
 QY 954 PKHFIDFCTMAECS 967







DR EMBL: AJ315733; CAC86014.1; --  
 DR HSP: P07996; ILSL.  
 DR MEROPS; M12.025; --  
 DR Genew; HGNC:16305; ADAMTS15.  
 DR MIM: 607509; --  
 DR InterPro; IPR010294; ADAM\_spacer1.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR001590; Peptidase M12B.  
 DR InterPro; IPR002870; Peptidase M12B N.  
 DR InterPro; IPR001818; Pept M10A\_M12B.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR000884; TSPI.  
 DR Pfam; PF05986; ADAM\_spacer1; 1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR Pfam; PF00090; TSP 1; 3.  
 DR SMART; SM00209; TSP1; 3.  
 DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS00346; CYSTEINE\_SWITCH; FALSE\_NEG.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50214; DISINTEGRIN\_2; FALSE\_NEG.  
 DR PROSITE; PS50092; TSP1; 3.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Extracellular matrix; Glycoprotein; Hydrolase; Metalloprotease;  
 KW Repeat; Signal; Zinc; Zymogen.  
 FT SIGNAL 1 17 Potential.  
 FT PROPEP 18 212 By similarity.  
 FT CHAIN 213 950 ADAMTS-15.  
 FT DOMAIN 213 427 Metalloprotease.  
 FT DOMAIN 428 515 Disintegrin-like.  
 FT DOMAIN 516 571 TSP type-1 1.  
 FT DOMAIN 572 700 Cys-rich.  
 FT DOMAIN 701 838 Spacer.  
 FT DOMAIN 839 895 TSP type-1 2.  
 FT DOMAIN 896 949 TSP type-1 3.  
 FT SITE 174 174 Cysteine switch (Potential).  
 FT METAL 361 361 Zinc (catalytic) (By similarity).  
 FT ACT\_SITE 362 362 By similarity.  
 FT METAL 365 365 Zinc (catalytic) (By similarity).  
 FT METAL 371 371 Zinc (catalytic) (By similarity).  
 FT CARBOHYD 141 141 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 591 591 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 623 623 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 679 679 N-linked (GlcNAc...) (Potential).  
 SQ SEQUENCE 950 AA; 103286 MW; 5DFBE18285CCCC3B CRC64;

Query Match 46.9%; Score 2477.5; DB 1; Length 950;  
 Best Local Similarity 48.9%; Pred. No. 4.2e-152;  
 Matches 485; Conservative 153; Mismatches 252; Indels 101; Gaps 24;

QY 36 LLLAAALAVSDALGRPEDEELVVP--ELE-----RVP---GHGTRLRLLHAF 81  
 Db 1 MLLGLITLAPAGRTAGGEPREVVPRLDPDINGRRYYWRGSDGQGLIFQITAF 60  
 QY 82 DQQLDLVPPDSSFLAPGFTLQNGRKSGSDTFLP-----ETDLACHFYSGTVNGDPSSA 136  
 Db 61 QEDFYHLTPDAQFLAPASTEHLG-----VPLQLTGGSSDLRCFFYSGVDNAEPDSF 114  
 QY 137 AALSICEGVGRGAFYLLGEAYFTQPLPAASERLATAAPGKPPAPLQFHLRRNRQGVGG 196  
 Db 115 AAVSLCGGLRGAFYGRGAEYVISPFPNAS---APAAQRNSQGA---HLLQ---RRGVPG 165  
 QY 197 TCGVDDDEPRPTGKAETDEDEGTGEDGPWS-----PQDPALGVGQP--TGTS 247  
 Db 166 PSG-----DPTSRC-----GVASG--WNPAIRLALDPYKRRAGFESSRRRS 207  
 QY 248 IRKKRFVSHRYVETMLVADQSMABFPHGSLKHYLLTFLPSVAARLYKHPISRNSVLVV 307  
 Db 208 GRAKRFVSIPIRYVELTVVADESMVKFHGADLEHYLLTLLATARLYRHPISILNPIVV 267  
 QY 308 KILVTHDEQKPEVTSNAALTIRNFCNQHNPPSDRDAEHYDTAILFTRODLCSQTC 367  
 Db 268 KVLRLDRDSGPKVTGNAALTIRNFCNQHNPPSDRDAEHYDTAILFTRODLCSQTC 327

RESULT 8  
 AT58 HUMAN  
 ID AT58 HUMAN STANDARD; PRT; 890 AA.  
 AC Q9UP79; Q9NZS0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
 DE with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2) (METH-  
 DE 8).  
 GN Name=ADAMTS8; Synonyms=METH2;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=99367466; PubMed=10438512; DOI=10.1074/jbc.274.33.23349;  
 RA Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S.,  
 RA Lombardo M., Iruela-Arispe M.L.;  
 RT "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new  
 RT family of proteins with angio-inhibitory activity.";  
 RL J. Biol. Chem. 274:23349-23357(1999).  
 RN [2]  
 RP SEQUENCE OF 195-440 FROM N.A.  
 RX MEDLINE=20079168; PubMed=10610729; DOI=10.1006/geno.1999.6014;

QY 368 DTGLMADYGTCDPSPSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACASLNGVND 427  
 Db 328 DTGLMADYGTCDPSPSCSVIEDDGLQAAFTTAHELGHVFNPHDNDVCEVFGKLRAN 387  
 QY 428 HMASMLNLSHDSQWSPSCSGYMITSLFDNGHGECIMDKPQNPPIQLPGLPSTSDANRQ 487  
 Db 388 HMASPTLIQIDRANPWSACSAIIITDLDGSHGDCLLDQPSKPISLPDLPGASYTLQQ 447  
 QY 488 COFTGEDSKHCPDAAASTCSTLWCTGTSGVLVCTKHFPWADGTSCEGKWCINGKCVN 547  
 Db 448 CELAFGVGSKPCP-YMQVCTKLWCTGKAGQWVCQTRHFPWADGTSCEGKCLKAGC 506  
 QY 548 K-NHRKHFDTPFHGSGWGMWPGDCSRTCCGGGVQVYTMRECDNPVPKNGKYGCEGRVYR 606  
 Db 507 RHNLNKH--RVDGSWAKWDPVGPCSRTCCGGGVQLARQCNTPTPANGKYGCEGRVYR 563  
 QY 607 SCNLEDPCDN-NGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKRCCLKICQAK 665  
 Db 564 SCNLEPCPSASGSKSFREEQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRAN 623  
 QY 666 GIGYFFVLQPKVVDGTPCSPDSTSVCGOQCVKAGCDRIIDSKKFKDKCGVCGGNGSTCK 725  
 Db 624 GTGFYVLAPKVVDTLCSPTSTSVCGOQCVKAGCDRIIDSKKFKDKCGVCGGNGSTCK 683  
 QY 726 KISGSVTSAPKPYHDIITPTGATNIEVQRNQRNNGSFLAIKAAADGTIYILNGDYTL 785  
 Db 684 KVTGLFTKPMGYNFVVAIPAGASSIDIRQGYKGLIGDDNYLALKNSQGYLLNGHFV 743  
 QY 786 STLEQDINMKGVLRYSGSSAALERIRSFPLKPLTIOVLTVGNALPKIYTYFVKK 845  
 Db 744 SAVERDLVVGSLRLYSQGTAVESLQASRPILPTVEVLVSQKMTTPRVRYSPYLPKE 803  
 QY 846 ---KES-----FNAIPTFS-----AWVEEWGECSCSCLGWQRRL 878  
 Db 804 PREDKSHPKDPRGSPVJHNSVLSUNVQEPDPPARWVAGSWGPCASCGSLQKRA 863  
 QY 879 VECRDINGQ---PASECAKEVKPASTRCPADHPQWOLGEMSSCKTCGKYKTKSLKC 935  
 Db 864 VDCRGSAGQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCKSKSGRFGRRSLKC 920  
 QY 936 LSHDGVLSHSDSCDPLKPKHPIDFCTWAE 966  
 Db 921 VGHGRLILARDQCNLHRKQPE-LDFCVLRPC 950

RA Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;  
 RT "ADAM-7S8, a novel metalloprotease of the ADAM-7S family located on  
 RT mouse chromosome 9 and human chromosome 11.";  
 RL Genomics 62:312-315(1999).  
 CC -!- FUNCTION: Has anti-angiogenic properties.  
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 CC matrix (By similarity).  
 CC -!- TISSUE SPECIFICITY: Highly expressed in adult and fetal lung,  
 CC lower expression in brain, placenta, heart, stomach and fetal  
 CC brain and kidney.  
 CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important  
 CC for a tight interaction with the extracellular matrix.  
 CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By  
 CC similarity).  
 CC -!- SIMILARITY: Belongs to the peptidase M12B family.  
 CC -!- SIMILARITY: Contains 1 disintegrin-like domain.  
 CC -!- SIMILARITY: Contains 2 TSP type-1 domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF060153; AAD48081.1; -;  
 DR EMBL; AF175283; AAF25806.1; -;  
 DR HSP; P07996; ILSL.  
 DR MEROPS; M12.226; -;  
 DR Genew; HGNC:224; ADAMTS8.  
 DR MIM; 605175; -;  
 DR GO; GO:0005178; F: integrin binding; TAS.  
 DR GO; GO:0009673; F: low affinity phosphate transporter activity; TAS.  
 DR GO; GO:0008237; F: metalloprotease activity; TAS.  
 DR GO; GO:0008285; P: negative regulation of cell proliferation; TAS.  
 DR InterPro; IPR010294; ADAM spacer1.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR006025; Pept\_M12B\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP 1.  
 DR Pfam; PF05986; ADAM\_spacer1; 1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR PROSITE; PS0215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS0214; DISINTEGRIN\_2; FALSE\_NEG.  
 DR PROSITE; PS0092; TSP1; 2.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Extracellular matrix; Glycoprotein; Heparin-binding; Hydrolase;  
 KW Metalloprotease; Repeat; Signal; Zinc; Zymogen.  
 FT SIGNAL 1 27 Potential.  
 FT PROPEP 28 214 By similarity.  
 FT CHAIN 215 890 ADAMTS-8.  
 FT DOMAIN 215 430 Metalloprotease.  
 FT DOMAIN 439 526 Disintegrin-like.  
 FT DOMAIN 527 582 TSP type-1 1.  
 FT DOMAIN 584 690 Cys-rich.  
 FT DOMAIN 691 832 Spacer.  
 FT DOMAIN 834 889 Poly-Pro.  
 FT DOMAIN 202 205 TSP type-1 2.  
 FT METAL 364 365 Zinc (catalytic) (By similarity).  
 FT ACT SITE 365 365 By similarity.  
 FT METAL 368 368 Zinc (catalytic) (By similarity).  
 FT METAL 374 374 Zinc (catalytic) (By similarity).  
 FT CARBOHYD 375 345 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 401 401 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 466 466 N-linked (GlcNAc...) (Potential).

FT	CARBOHYD	491	491	N-linked (GlcNAc...)	(Potential).
FT	CARBOHYD	600	600	N-linked (GlcNAc...)	(Potential).
FT	CONFLICT	195	195	E -> R (in Ref. 2).	
FT	CONFLICT	413	440	YLTLLDGGHGDCLLDARGAALPLPTGL -> FSCSHLQGW	
FT	SEQUENCE	890	AA; 96671	MW; 57D70EE03D5739D3	CRG64;
FT	SEQUENCE	890	AA; 96671	MW; 57D70EE03D5739D3	CRG64;
Query Match		43.9%	Score 2320.5;	DB 1;	Length 890;
Best Local Similarity		49.5%	Pred. No. 6e-142;		
Matches		450;	Conservative 147;	Mismatches 250;	Indels 63;
					Gaps 22;
QY	33	VPTLLLAALLAVS-DALGRPSR--EDELVPELERVPGH-GTTRLRLHAFPOQLDLD	88		
DB	11	LPFLLLLLLPLARGAPARPAAGGQASELVVP--TRLPGSAGELALHLASFAFGKGVLR	68		
QY	89	VPPSSFLAPGFTLQNV---GRKSGSDTLPETDLAHCFFYSGTVNGDPSSAAALSCEGV	145		
DB	69	LAPDSSFLAPEFKIERLGGSGRATGG-----ERLGRGCFSGTVNGDPSSAAALSCEGV	123		
QY	146	RGAYLLGEAYFIQPLPAAS-----ERLATAAPCEKPPAPLQFHLLRRNRQDVGTCGV	200		
DB	124	SGSFLDGEFTIQPGAGGSLAQPHRLRWGPAGARLP-----RGPEWEV-----ET	172		
QY	201	VDEPRPTGKAETDEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGE	260		
DB	173	GEGQQRGERQHDQESDESE	222		
QY	261	ETMLVADQSMAEFHSGGLKHYLLTLFVSAARLYKHPSIRNSVSVLVVVKILVHDEQKPE	320		
DB	223	ETLLVADQSMAEFHSGGLKHYLLTLFVSAARLYKHPSIRNSVSVLVVVKILVHDEQKPE	282		
QY	321	VTSNAALTNRNFCNWKQKHNPPSDRDAEHVDYDTALLFTRQDLCGSGQ--TCDTLGMADV	379		
DB	283	VSDNGGLTLNRNFCNWKQKHNPPSDRDAEHVDYDTALLFTRQDLCGSGQ--TCDTLGMADV	342		
QY	380	DPSSCSVIDEDGLOAAFTTAHELGHVFNMPHDDAKOCASLNGVQDQSHMMASMLNDH	439		
DB	343	DPNSCSVIDEDGLOAAFTTAHELGHVFNMPHDDAKOCASLNGVQDQSHMMASMLNDH	402		
QY	440	SQWSPGSGYMITSLFNGHGECLMDKPNQIOLPGDLPG--TSYDANROCOFTFGEDSK	497		
DB	403	TLWSPGSGYMITSLFNGHGECLMDKPNQIOLPGDLPG--TSYDANROCOFTFGEDSK	462		
QY	498	HCPD--AASTCTLWCTGTSGVLVCOQTKH--FPWADGTCGEGKWCINGKCNVKNHRKH	553		
DB	463	HCPNTSAQDVCAQLWC-HTDGAEPCHTKNGSLPWADGTPCGPHLCSGSCLPPEEVER	521		
QY	554	FDPFHSGWGMWPGWDCSRCTCGGVQYVTRWCNDPVPKNGGKYCEGKRVYRSCNLEDC	613		
DB	522	PKPVVDGWAAPWPGWGCSTRTCGGVQFVSHRECKDPBPQNGRYCLGRRAKYQSCHTEEC	581		
QY	614	PDNNGKTFREOCEAHNEFSAFSGSPAVEMIPKAGVSPKDRCKLICQAKGIGYFPVL	673		
DB	582	PP-DGKSFREOCEAHNEFSAFSGSPAVEMIPKAGVSPKDRCKLICQAKGIGYFPVL	639		
QY	674	QPKVVDGTPGSPDSTSVCOQCCKVAGCDRIIDSKKFKDKCGVCGNGSGTCKKISGVS	733		
DB	640	EAKVIDGLCGPETALICVRGQCVKAGCDHVDSPRKLDCGVCCKGNSCRKVSGLTP	699		
QY	734	AKPYHDIITPTGATNIEVKQRNQRNGSGFLAKAADGTILNGDYTLSTLEQDIM	793		
DB	700	TNYGNDIVTIPAGATNIDVKQRSHPGVQNDGNYLAKLTADQYLLNGNLNLAISAIEQDIL	759		
QY	794	YGVVLYRSGSSAALRIRSPPLKPLTIQVLTIV-GNALRPKIKYTFV-----842			
DB	760	VKGILKYSIATLERLQSFRLPPELTVPGVFPFKYTFVFDVNDVDFSMQS	819		
QY	843	KKKESFNAI--PTESA-WVIEEWGECSSCELGQRRLVECRDINGQFASCAKEVPAS	900		
DB	820	SKERATTNIIQPLLAQWLVLDWSECSCTCGAGQWRRTVECRDPSGQASATCNKALKPED	879		
QY	901	TRPCADHPCP	910		

Db 880 AKPCSQLCP 889

RESULT 9

ATS8 MOUSE

ID ATS8 MOUSE STANDARD; PRT; 905 AA.

AC P57110;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2).

GN Name=Adamts8;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20079168; PubMed=10610729; DOI=10.1006/geno.1999.6014;

RA Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;

RT "ADAM-TS8, a novel metalloproteinase of the ADAM-TS family located on mouse chromosome 9 and human chromosome 11.";

RL Genomics 62:312-315(1999).

CC -!- FUNCTION: Has anti-angiogenic properties (By similarity).

CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).

CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).

CC -!- TISSUE SPECIFICITY: Expressed specifically in adult lung and heart and low expression during mouse development.

CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix.

CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By similarity).

CC -!- SIMILARITY: Belongs to the peptidase M128 family.

CC -!- SIMILARITY: Contains 1 disintegrin-like domain.

CC -!- SIMILARITY: Contains 2 TSP type-1 domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AF175282; AAF25805.1; -;

DR HSPF; P07996; 1LSL.

DR MEROPS; M12.226; -.

DR MGD; MGI:1353468; Adamts8.

DR InterPro; IPR010294; ADAM spacer1.

DR InterPro; IPR001762; Disintegrin.

DR InterPro; IPR006025; Pept\_M\_Zn\_BS.

DR InterPro; IPR001590; Peptidase\_M128.

DR InterPro; IPR000884; TSP1.

DR InterPro; IPR008085; TSP 1.

DR Pfam; PF05986; ADAM spacer1; 1.

DR Pfam; PF01421; Reprolysin; 1.

DR Pfam; PF00090; TSP 1; 2.

DR PRINTS; PRO1705; TSP1REPEAT.

DR SMART; SM00209; TSP1; 2.

DR PROSITE; PS50215; ADAM\_MPRO; 1.

DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE NEG.

DR PROSITE; PS50214; DISINTEGRIN\_2; FALSE\_NEG.

DR PROSITE; PS50092; TSP1; 2.

DR PROSITE; PS00142; ZINC\_PROTEASE; 1.

KW Extracellular matrix; Glycoprotein; Heparin-binding; Hydrolase;

KW Metalloprotease; Repeat; Signal; Zinc; Zymogen.

FT SIGNAL 1 28 Potential.

FT PROPEP 29 228 By similarity.

FT CHAIN 229 905 ADAMTS-8.

FT DOMAIN 229 452 Metalloprotease.

FT DOMAIN 453 541 Disintegrin-like.

FT	DOMAIN	542	597	TSP type-1 1.
FT	DOMAIN	599	705	Cys-rich.
FT	DOMAIN	706	847	Spacer.
FT	DOMAIN	848	904	TSP type-1 2.
FT	METAL	378	378	Zinc (catalytic) (By similarity).
FT	ACT_SITE	379	379	By similarity.
FT	METAL	382	382	Zinc (catalytic) (By similarity).
FT	METAL	388	388	Zinc (catalytic) (By similarity).
FT	CARBOHYD	415	415	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	480	480	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	506	506	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	615	615	N-linked (GlcNAc. . .) (Potential).
SQ	SEQUENCE	905 AA;	98879 MW;	124D4132B33A0CAE CRC64;

Query Match 42.9%; Score 2265.5; DB 1; Length 905;

Best Local Similarity 47.9%; Pred. No. 2.3e-138;

Matches 441; Conservative 146; Mismatches 263; Indels 71; Gaps 18;

QY	34	PTLLLLAAALLAVSDALGRPS-----BEDELVVPELSEVPEHGT-TLRLHAFDQQLDL	87
Db	11	PPLLLLLLQPLPPPLVCGAPAGPGTGAQASELVVP--TRLFGSASELAFHLSAFQGGFVL	68
QY	88	DVPDSSFLAPGFTLQNVGRKSGSDTLPETDLAHCFSYGVTVNGDPPSSAAALSCEGVRG	147
Db	69	RLAPDASFLAPEFKIERLGGSSAAG--GEPLRGCFSGIVNGERESLAAMSCVAGWSG	126
QY	148	AFYLLGEAYFTQPLPAAS-----ERLATAAPGKXP-----PAPLQFHLLRN	189
Db	127	SFLAGEEFTIQPGAGDLSLQPHLRQWPGQRREDPLAAAEVFPPLPQGLEWEVENGN	186
QY	190	RQDVGTCGVVDDEPRPTGKAETEDDEGTGEGDEGPQWSPQDPALGVQPTGTGSIR	249
Db	187	QGGQ-----ERSDNEEDRKQKEGLLKETEDS--RKVPPTFGS-KTR	225
QY	250	KKRFVSSHRYVETMLVADQSMAPFHGSLKHYLLTFSVAARLYKHPISRVSLVVKI	309
Db	226	SKRFVSEARFVETLLVADASMAAFYGLDLQNHITVMSMAARIYKHPISRVSLVVKV	285
QY	310	LVIHDEQKGPVTSNAALTLENFQWQHPSPDRDAEHYDTAILFTFQDLQG-SQTC	368
Db	286	LIVEKRWGPVSDNGGLTLRFNFCWQRFRNKPSPDRHPEHYDTAILFTFQDLQG	345
QY	369	TLGMADVTGCDPRSRSVIEDDGLQAAFTTAHELGHVFNPHDDAKOCASLNGVQDSH	428
Db	346	TLGMADVTGICDPDKSCSVIKDEGLQAAFTTAHELGHVLSMPHDDSKPCVRLFGPMGKYH	405
QY	429	MWASMLSNLDHSQWSPSCGYMITSFLDNGHGECLMDKPNPIQLPGDLPG--TSYDANR	486
Db	406	MMAPFFIHVNTLPWSPCSAVYLTLLDDGHGDCCLDAPTSVLPPLPGLPGHSTLYELDQ	465
QY	487	QCQTFEGEDSKHCPDAA--STCSTLWCTGTSGGVLVCQTKH--FPWADGTSCEGKWCIN	542
Db	466	QCKQIFGPDPRHCPNTSVEDICVQLCARHRSDEPICHTKNGSLWADGTFCGPGHLCLD	525
QY	543	GKCVNKNHRKHFDTPFHGSGWGWPGDCSRTCCGGVQVYTWRECDNPVKNKGKCYCEGR	602
Db	526	GSCVLKEDVENPKAVVDGDMGFWPWCQSCRTCCGGIQFSNRECDNPMQNGRFLGGR	585
QY	603	VRYBSCLNDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWPKYAGVSPKDRCKLIC	662
Db	586	VKYQSCNTEECPP--NGKSFREQCEKYNAYNHTDL-DGNFLQWPKYSGVSPDRCKLFC	643
QY	663	QAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQCQKAGCDRIIDSKKFKDCGCGGNGS	722
Db	644	RARGSEFKVFEAKVIDGTLCGPDTLSCVVRGQCVKAGCDHVNSPKLKDCKGCGGKGT	703
QY	723	TCKKISSVTSAPKPGYHDIITPTGATNIEVKQNRQGRNNGSFLAKAADGTYILNGD	782
Db	704	ACRKISGSFTPFYSYGYNDIVTIPAGATNIDVKQRSHPGVRNDGSLALKTANGQYLLNGN	763
QY	783	YTLSTLEODIMYKGVWLYRSGSSAALERSFSPLKPLTIOVLTV-GNALRPKIKTYTF	841
Db	764	LAISAIEDIILVKGTILKYSGMATLERLQSFQALPEPLTVQLLTVSGSEVPPKRYTYFF	823



```
Db 37 LLLLLLALLSARLASPLRBEIEIVFPEKLNGSVLPGSGAPARLLCRLOAFGETLLLEL 96
QY 90 PPDSSFLAPGFTLQNVGRK----SGSDTLPETDLAHCFYSGTVNGDPSSAAALSCEG- 144
Db 97 EQDSGVQVGLTVQYLQAPPELLGAE---PGT-----YLTGTINGDPESVASLHWDGA 148
QY 145 VNGAFYLLGEAFYIQLPAAASERLATAAAGEKPPAPLQPHLLRRNRQGVGTGCVVDDE 204
Db 149 LLGVLYQYGAELHQLPLEGGTNSA--GGPGA-----HILRRK----- 184
QY 205 PRPTGKAETDEDEGTEGEDEGPQSPDPAALQGVQPTGTGTSIRKKRVSSHRYVETML 264
Db 185 -----SPASGQGPNCVKAP-----LGSPSPRPR--RAKRFASLSRFPVETLV 224
QY 265 VADQSMAPFHGSLKHYLLTLFSAARLYKHPSIRNSVSLVVVKILVIHDEQKGPVTSN 324
Db 225 VADDRKAAAFHAGLKYLLTVMAAAAKAFKHPISIRNPVSLVTRLVILGSGEGVQVGPS 284
QY 325 AALTURNFCNWKQKHNPPSDRAEHYDTAILFTRODLGSGQTCDTLGMADVTVCDPSRS 384
Db 285 AAQTLRSFCAMQRLNTPEDSDPDHFDTAILFTRODLGCVSTCDTLGMADVTVCDPARS 344
QY 385 CSVIEDDGLQAFTTAHELGHVFNPHDDAKOCASLNG--VNQDSHMMASMLNLDHSPQW 443
Db 345 CAIVEDDGLQSAFTAHELGHVFNMLHDNSKPCISLNGPLSTRHVMAPVMAHVDPPEPW 404
QY 444 SPCSGYMITSLFDNGHGECLMDKQNPQIQLPGDIPGTSYDANROQOFTFGEDSKHCPDAA 503
Db 405 SPCSGARFITDFLDNGYGHCLLDKPEAPLHLPVTFPGKYDADROQLTFGPDPSRHCPLP 464
QY 504 STCSTLWCTGTSGGVLVVQTKHFPWADGTSCEGKWCINGKCVNKNRKHFDTPFHGSGW 563
Db 465 PPCAALWCSGHNLGHAMCQTKHSPWADGTPCPGAQACMGGRCLHMDQLQDFNIPQAGGWG 524
QY 564 MWGPWGDCSRTCGGVQVYVNRCDNPVPKNGKCYCEKRVYRSCNLEDCPDNNNGKTRE 623
Db 525 PWGPWGDCSRTCGGVQVFSRDCRTPVRPNNGKYCEGRTRFRSCNTEDCPTGSALTFRE 584
QY 624 ECEAHNBFKASFGSGPA--VWIPKYAGVSPKDRCKLTCQAKGIGYRPVLPKVVVDGTP 682
Db 585 EQCAAYNHRDIL--FKSFPQMDVRYTGVAFQDQCKLTCQARALGYIYVLEPRVVDGTP 643
QY 683 CSPDSTSCVQCQVKAGCDRIIDSKKFKDKCGVCGNGSGTCKISGSVTSAKPGYHDII 742
Db 644 CSPDSSVCVQCRCIHAGCDRIIGSKKFKDKMVCGGDGGSGCKSGSPRKYGVNVV 703
QY 743 TTPGTATNIEVQRNQRNGRNSFLAIIKADGTIYLNQDYTLSTLEQDQIMYKGVV--LRY 801
Db 704 TTPAGATHILVRQQGNPHGRS--IYALKLDPGYSALNGEYTLMPSPDTDVILFGAVSLRY 761
QY 802 SGSSAALERIRSFSPLEPLTIQVLTGVNALRPKIKYTVFVKKKESFNAIPTFSAWV 859
Db 762 SGATAASELTSGHGGLAQLPTLIQVLVAGNPQDTRLRYSFVPRPTFS--TPRPTQDWL 818

RESULT 11
Q6UWA8 PRELIMINARY; PRT; 837 AA.
AC Q6UWA8;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE ADAMTS4.
GN ORFNames=UNQ769;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
```

```
RA Eaton D., Foster J., Grimaldi C., Gu Q., Haas P.E., Heldens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R., Watanabe C., Wieand D., Woods K., Xie M.H., Yansura D.,
RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
RA Godowski P.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RL Genome Res. 13:2265-2270(2003).
DR EMBL; AY36886; AAQ9245.1; -.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR010294; ADAM_spacer1.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR006025; Pept_M_zh_BS.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR Pfam; PF05986; ADAM_spacer1; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR PRINTS; PR01705; TSP1; 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS02015; ADAM_MEPRO; 1.
DR PROSITE; PS00092; TSP1_1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 1.
SQ SEQUENCE 837 AA; 90166 MW; 5DF9C9ACF67CF8F CRC64;

Query Match 40.0%; Score 2117; DB 2; Length 837;
Best Local Similarity 48.9%; Pred. No. 8.9e-129;
Matches 410; Conservative 132; Mismatches 226; Indels 70; Gaps 18;

QY 36 LLLLLAALLAVSDALGRPSEDEELVPEL---BRVPGHT--TRL--RLHAFDQQLDLDV 89
Db 37 LLLLLLALLSARLASPLRBEIEIVFPEKLNGSVLPGSGAPARLLCRLOAFGETLLLEL 96
QY 90 PPDSSFLAPGFTLQNVGRK----SGSDTLPETDLAHCFYSGTVNGDPSSAAALSCEG- 144
Db 97 EQDSGVQVGLTVQYLQAPPELLGAE---PGT-----YLTGTINGDPESVASLHWDGA 148
QY 145 VNGAFYLLGEAFYIQLPAAASERLATAAAGEKPPAPLQPHLLRRNRQGVGTGCVVDDE 204
Db 149 LLGVLYQYGAELHQLPLEGGTNSA--GGPGA-----HILRRK----- 184
QY 205 PRPTGKAETDEDEGTEGEDEGPQSPDPAALQGVQPTGTGTSIRKKRVSSHRYVETML 264
Db 185 -----SPASGQGPNCVKAP-----LGSPSPRPR--RAKRFASLSRFPVETLV 224
QY 265 VADQSMAPFHGSLKHYLLTLFSAARLYKHPSIRNSVSLVVVKILVIHDEQKGPVTSN 324
Db 225 VADDRKAAAFHAGLKYLLTVMAAAAKAFKHPISIRNPVSLVTRLVILGSGEGVQVGPS 284
QY 325 AALTURNFCNWKQKHNPPSDRAEHYDTAILFTRODLGSGQTCDTLGMADVTVCDPSRS 384
Db 285 AAQTLRSFCAMQRLNTPEDSDPDHFDTAILFTRODLGCVSTCDTLGMADVTVCDPARS 344
QY 385 CSVIEDDGLQAFTTAHELGHVFNPHDDAKOCASLNG--VNQDSHMMASMLNLDHSPQW 443
Db 345 CAIVEDDGLQSAFTAHELGHVFNMLHDNSKPCISLNGPLSTRHVMAPVMAHVDPPEPW 404
QY 444 SPCSGYMITSLFDNGHGECLMDKQNPQIQLPGDIPGTSYDANROQOFTFGEDSKHCPDAA 503
Db 405 SPCSGARFITDFLDNGYGHCLLDKPEAPLHLPVTFPGKYDADROQLTFGPDPSRHCPLP 464
QY 504 STCSTLWCTGTSGGVLVVQTKHFPWADGTSCEGKWCINGKCVNKNRKHFDTPFHGSGW 563
Db 465 PPCAALWCSGHNLGHAMCQTKHSPWADGTPCPGAQACMGGRCLHMDQLQDFNIPQAGGWG 524
QY 564 MWGPWGDCSRTCGGVQVYVNRCDNPVPKNGKCYCEKRVYRSCNLEDCPDNNNGKTRE 623
Db 525 PWGPWGDCSRTCGGVQVFSRDCRTPVRPNNGKYCEGRTRFRSCNTEDCPTGSALTFRE 584
QY 624 ECEAHNBFKASFGSGPA--VWIPKYAGVSPKDRCKLTCQAKGIGYRPVLPKVVVDGTP 682
Db 585 EQCAAYNHRDIL--FKSFPQMDVRYTGVAFQDQCKLTCQARALGYIYVLEPRVVDGTP 643
QY 683 CSPDSTSCVQCQVKAGCDRIIDSKKFKDKCGVCGNGSGTCKISGSVTSAKPGYHDII 742
Db 644 CSPDSSVCVQCRCIHAGCDRIIGSKKFKDKMVCGGDGGSGCKSGSPRKYGVNVV 703
QY 743 TTPGTATNIEVQRNQRNGRNSFLAIIKADGTIYLNQDYTLSTLEQDQIMYKGVV--LRY 801
Db 704 TTPAGATHILVRQQGNPHGRS--IYALKLDPGYSALNGEYTLMPSPDTDVILFGAVSLRY 761
QY 802 SGSSAALERIRSFSPLEPLTIQVLTGVNALRPKIKYTVFVKKKESFNAIPTFSAWV 859
Db 762 SGATAASELTSGHGGLAQLPTLIQVLVAGNPQDTRLRYSFVPRPTFS--TPRPTQDWL 818
```

Db 525 PWGPMGDCSRTCGGVQVQFSSRDTRPVPRNGKGYCEGRRTRFRSCNTEDCPTGTSALTFR 584  
QY 624 EQCEAHNEFSKASRSGSPA-VWIPKYAGVSPKDRCKLIKQAKGIGYFVLPKPVVDGTP 682  
Db 585 EQCAAYNHRTDL-FKSPFGPMQWVRYTGAPQDQCKLTQCARALGYVYVLEPRVVDGTP 643  
QY 683 CSPDSTSVCGQCQVKAGCDRIIDSKKPKDKGVCGNGSTCKKISGVSPTSAPKPGYHDI 742  
Db 644 CSPDSSVCVQGRCHTAGCDRIIGSKKPKDKWCVGGDGGSGSKSGSFRKFRYGYNNV 703  
QY 743 TIPTGATNIEVKQRNQRNNGSFLATKAADGTYYLNGDYTLSTLEODIMYKGVV-LRY 801  
Db 704 TIPAGATHILVRQGNPGHRS--IYLALKLPDGSYALNGEYTLMPSTDWLPGAVSLRY 761  
QY 802 SSSAALERIRSFSPKLEPLTIOVLTVGNALRPKIKYFYVKKKESFNAPITPFSAW 859  
Db 762 SGATAASETLSHGFLAQLPLTLQVLVAGNPQDTRLRYSFVPRPTPS-TPRPTPDWL 818

RESULT 12  
Q6P4Q8 PRELIMINARY; PRT; 837 AA.  
ID Q6P4Q8  
AC Q6P4Q8  
DT 05-JUL-2004 (TremBLrel. 27, Created)  
DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TremBLrel. 27, Last annotation update)  
DE A disintegrin and metalloproteinase with thrombospondin motifs-4,.  
DE 4,.  
QN Name=ADAMTS4;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulláhy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smalius D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Strausberg R.;  
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC063293; AAH63293.1; -  
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
DR GO; GO:0007229; P:proteolysis-mediated signaling pathway; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR005886; ADAM cysteine.  
DR InterPro; IPR010294; ADAM spacer1.  
DR InterPro; IPR001590; Peptidase M12B.  
DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR008085; TSP.1.  
DR Pfam; PF05986; ADAM\_spacer1; 1.

DR Pfam; PF01421; Reprolysin; 1.  
DR Pfam; PF00090; TSP.1; 1.  
DR PRINTS; PR01705; TSP1REPEAT.  
DR SMART; SM00608; ACR; 1.  
DR SMART; SM00209; TSP1; 1.  
DR PROSITE; PS02015; ADAM MEPRO; 1.  
DR PROSITE; PS00092; TSP1; 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
KW Integrin.  
SQ SEQUENCE 837 AA; 90149 MW; 3B91C651E54EFC5F CRC64;  
Query Match 39.8%; Score 2106; DB 2; Length 837;  
Best Local Similarity 48.7%; Pred. No. 4.6e-128;  
Matches 408; Conservative 134; Mismatches 226; Indels 70; Gaps 18;  
QY 36 LLLLAALLAVSDALGRSPSEDEELVPEL---BRVPGHGT-TRL--RLHAFDQOLDLDV 89  
Db 37 LLLLLLASLPASRLASPLPREEBIVFPEKLNGSVLPSCGAPARLLCRLQAFGETLLIEL 96  
QY 90 PPDSSFLAPGFTLQNVGRK---SGSDTPLPETDLAHCIFYSGTVNGDPSSAAALSCEG- 144  
Db 97 EQDSGVQVEGLTVQYLQAPELGGAB---PGT-----YLTGTINGDPESVASLHWGG 148  
QY 145 VRGAFYLLGEAYFTIPLPAASERLATAPGKPPAPLQFHLRRNRQDVGSGTCGVVDDE 204  
Db 149 LLGVLYRGAELHLQPLEGGTPNSA-GPGCA-----HILREK----- 184  
QY 205 PRPTGKAETDEDEGTGEDEGPQSPDPAQGVQPTGTSIRKIRKRVSSHRYVETML 264  
Db 185 -----SPASGGPMCNVKAP-----LGSPSPRPR-RAKRFASLSRFVETLV 224  
QY 265 VADQSMAEFFHGSGIKHYLLTLFSVAARLYKHPSTRNSVLSLVVVKILVIHDEQKPEVTSN 324  
Db 225 VADKMAAFRFGAGLRYLLTVMAAAAKAFKHPSTRNPVSLVTVKLVILGSGEGPQGVPS 284  
QY 325 AALTNRNFCNWKQHNPPSDRDAEHYDTAILTRQDLCSGQTCDTLGMADVGTVCDFPSRS 384  
Db 285 AAQTLRSFCAWQRLNTPENSDPDHPDTAILTRQDLCGVSTCDTLGMADVGTVCDFPSRS 344  
QY 385 CSVIEDGLQAAFTTAHELGHVFNMPHDDAKQACASLNG-VNDSHMMASMLNLDHSQPV 443  
Db 345 CAIVEDGLQSAFTAAHELGHVFNVLHNSKPCISLNGPLSTSRVNVNAPVMAHVDPEPW 404  
QY 444 SPGSGYMITSLDNGHCEGLMDKPONPIQLPGDLPGTSYDANRCQFTFCEDSKHCDDAA 503  
Db 405 SPGSAFITDNLNGYGHCLLDKPEAPLHPVTFPGKDYADRCQQLTFGPDSCRHCPQLP 464  
QY 504 STCSLWCTGTSGVLVCQTQKHPFPAWADTSCGEGKWCINGKCNKHNKHFDTFPFGHSWG 563  
Db 465 PPCAALWCSGLHNGHAMCQTKHSPWADGTPCGPAQACMGGRCLHMDQLQDFNIPAGGWG 524  
QY 564 MWGPMGDCSRTCGGVQVYTWRECDNPVRNGKGYCEGKRVRYRSCNLEDCPDNNKGTFR 623  
Db 525 PWGPMGDCSRTCGGVQVQFSSRDCTRPVTRNGKGYCEGRRARFRSCNTEDCPTGSAITFR 584  
QY 624 EQCEAHNEFSKASRSGSPA-VWIPKYAGVSPKDRCKLIKQAKGIGYFVLPKPVVDGTP 682  
Db 585 EQCAAYNHRTDL-FKSPFGPMQWVRYTGAPQDQCKLTQCARALGYVYVLEPRVVDGTP 643  
QY 683 CSPDSTSVCGQCQVKAGCDRIIDSKKPKDKGVCGNGSTCKKISGVSPTSAPKPGYHDI 742  
Db 644 CSPDSSVCVQGRCHTAGCDRIIGSKKPKDKWCVGGDGGSGSKSGSFRKFRYGYNNV 703  
QY 743 TIPTGATNIEVKQRNQRNNGSFLATKAADGTYYLNGDYTLSTLEODIMYKGVV-LRY 801  
Db 704 TIPAGATHILVRQGNPGHRS--IYLALKLPDGSYALNGEYTLMPSTDWLPGAVSLRY 761  
QY 802 SSSAALERIRSFSPKLEPLTIOVLTVGNALRPKIKYFYVKKKESFNAPITPFSAW 859  
Db 762 SGATAASETLSHGFLAQLPLTLQVLVAGNPQDTRLRYSFVPRPTPS-TPRPTPDWL 818

RESULT 13

```
Q8K384
ID Q8K384 PRELIMINARY; PRT; 833 AA.
AC Q8K384;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Name=Adamts4;
GN Adamts4 protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullishy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
SEQUENCE FROM N.A.
RP STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027773; AAH27773.1; -.
DR HSSP; P07996; 1LSL;
DR MEROPS; M12.221; -.
DR MGD; MGI:133949; Adamts4.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR010294; ADAM_spacer1.
DR InterPro; IPR001590; Peptidase M12B.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP1.
DR Pfam; PF05986; ADAM_spacer1; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; TSP1; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKOWN 1.
SQ SEQUENCE 833 AA; 90097 MW; 4CEC83DFBC3AA619 CRC64;

Query Match 39.8%; Score 2105; DB 2; Length 833;
Best Local Similarity 48.7%; Pred. No. 5.3e-128;
Matches 409; Conservative 131; Mismatches 221; Indels 78; Gaps 19;

QY 28 RSFGVPVTL-----LLLAALLAVDALGRPSDEELVVE-----LERVGHGT-TR 75
DB 18 RRFQPCLPHTVQWRRLLLAFLLSLAWP-ASPLPREEBIVFPEKINGSSILPGSGVPAR 76
QY 76 L--RLHAFDQDLQDVPDPSSFLAGFTLQNGRK-----SGSDTLPETDLAHCFYSGTV 129
DB 77 LLYRLPAFGEMLLLEQDPGVQVEGLTVQYLGAPEMLGGAE---PGT-----YLTGTI 128

130 NGDPSSAAALSCEG-VGAFYLLGEAYFIQPLPAASERLATAAPGEKPPAPLOPHLLRR 188
129 NGDPESVASLHWGSGALLGVLYRGAELHLQPLEGALNSA-GGPGA-----HILRR 179
189 NRQGDVGTCTGVVDDEPRPTKAEDEDEGTEDEGEPQSPQDPALQGVQOPTGTGSI 248
180 K-----SPASSQGMCTVKAPS-----GSPSPI-SR 204
249 RKKRFSVSHRYVETMLVADQSMABFPHGSLKHYLLTLTFSVAARLYKHPISIRNSVLVVVK 308
205 RTKPFASLSRFVETLVVADDDKAAAFHGTGLKRYLLTVMAAAAKAFKHSIRPNVNLVTR 264
309 ILVIHQEQKEPVTSSNAALTILRNFCNMQKQNPSPDRDAEHYDTAILTRODLCSQTC 368
265 LVILGSGEGEPQVGPSSAAQTLRSFCTWQGLTNPDSDPDHFDTAILTRODLCSVTC 324
369 TLGWADVCTVCDPSRSCSVIEDDGLQAFTTAHGLGHVFNPHDDAKOCASINGVQDS- 427
325 TLGWADVCTVCDPARSCAIVEDDGLQSAFTAHAHGLGHVFNHLDNSKPTNLNGOGSSR 384
428 HMAASMLNLDHSPWSPSCGYMITSFLDNGHGECLMDKPNQPIQLPGDLPTGTSYDANRQ 487
385 HVMAFVMAHVDPPEPWSFCSARFITDFLDNGYGHCLLDKPEAPLHLPATFFPKDYADRQ 444
488 CQFTGEDSKHCPDAASTCSTLWCTGSGGVLVCTKHFPWADGTSCEGKWCINGKCVN 547
445 CQLTFGPDSSHCPLPPPCAALWCGLNHGAMCQTKHSPWADGTPCGSSQACMGGRCLH 504
548 KNRKHGFTDPHSGNMGWPGDCSRCTCGGCVGVVYTMRECDNPVKNKGKCYGEGKVRYS 607
505 VDQLKDFNVPAQGMGMPMGDCSRCTCGGCVGVVYSSRDCTRPVPRNGGKCYGEGKTRFRS 564
608 CNLEDCPNNGKTFREEOCEAHNEFSKASFGSGPA-VIEWPKYAGVSPKDKLITCOAKG 666
565 CNTENCPHGSALTFREEOCAAYNHRTDL-FKSPFQPMQVPRYTGVAARDCKLTCQARA 623
667 IGYFPVLQPKVVDGTPCSPDSTSVCVQGVKAGCDRIIDSKKPKDCGCVGCGNGSTCKK 726
624 LGYVYVLEPRVADGTPCSPDSTSVCVQGVKAGCDRIIDSKKPKDCGCVGCGNGSRCSK 683
727 IGSVTSKAPGYHDIITITGCTNIEVKQNRNORNGSFLAIAKADGTGVLNGDYTLS 786
684 QSGSFKKFRYGYSDVVTIPAGATHILVRQ--QGGSGLSKIYLAALKSDGSVANGYEYTL 741
787 TLEODIMYKGVV-LRYSGSSAALERSFSFLKPEPLTITQVLTGVNALRPKIKYTVFVKK 844
742 PSPTDVLPGVSLRYSGATAASETLSHGFLAQPLTLOVLVAGNPQVARNLRYSEFFVR 800
RESULT 14
Q8BNJ2 PRELIMINARY; PRT; 845 AA.
AC Q8BNJ2;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Mus musculus 9 days embryo whole body cDNA, RIKEN full-length enriched
DE library, clone: D030041M02 product: a disintegrin-like and
DE metalloprotease (reprolysin type) with thrombospondin type 1 motif, 4,
DE full insert sequence.
GN Name=Adamts4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Whole body;
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999).
```



RN  
 RP  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN  
 RP  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN  
 RP  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN  
 RP  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RA "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN  
 RP  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK083534; BAC38944.1; -.  
 DR HSSP; P07996; ILSL.  
 DR MEOPS; M12.221; -.  
 DR MGD; MGI:1339849; Adante4.  
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
 DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR010294; ADAM spacer1.  
 DR InterPro; IPR001590; Peptidase M12B.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP 1.  
 DR Pfam; PF05986; ADAM\_spacer1; 1.  
 DR Pfam; PF01421; ReprQ1ysin; 1.  
 DR Pfam; PF00090; TSP 1; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00209; TSP1; 1.  
 DR PROSITE; PS00215; ADAM\_MPRO; 1.  
 DR PROSITE; PS00092; TSP1; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN\_1.

KW Integrin; Metalloprotease; Protease.  
 SQ SEQUENCE 845 AA; 91238 MW; B496C3190D1A9225 CRC64;  
 Query Match 39.8%; Score 2104.5; DB 2; Length 845;  
 Best Local Similarity 49.2%; Pred. No. 5.9e-128;  
 Matches 405; Conservative 131; Mismatches 216; Indels 71; Gaps 18;  
 QY 37 LLLAAALLAVSDALGRPSEDEELVPE---LRRVPHGT-TRL--RLHAFDQDLDDV 89  
 DB 46 LLLAFLLSLAMP-ASPLPREEEIVFPEKLNGSSILFOSGVPARLLYRLPAPFGEMLLLEL 104  
 QY 90 PPDSSFLAPGFTLQNVGRK-----SGSDTLPETDLAHCFTSGTVNGVPSSAAALSCLBG- 144  
 DB 105 EQDFGVQVEGLTVQYLGAPEMLGAB---PGT-----YLTGTINGDPESVASLHWDGGA 156  
 QY 145 VRGAFYLLGEAYFTIPLPAASERLATAAPGKPPAPLOFHLLRRNRQDVGCGTCGVDD 204  
 DB 157 LLGLVQYRGAEFLHLQPLEGGALNSA-GPGGA-----HILRRK----- 192  
 QY 205 PRPTGKAETEDDESGTEGDEGPQWSPQDPALQGVGPTGTGSTRKRFVSSHRVYETML 264  
 DB 193 -----SPASSQGPMTVKAP-----GSFPI-SRTRKRFASLSRFVETLV 232  
 QY 265 VADQSMAEFFHSGGLKHYLLTLFSAARLYKHPISIRNSVLSLVVVKILVIHDEKQGPVTSN 324  
 DB 233 VADDKMAAFHGTGLKRYLLTVMAAAAKAFKHPISIRNPVNLVTVLTVILGSGQEGPQVGPS 292  
 QY 325 AALTIRNFCWQKHNPSPSDAHDYDTALTFTQDLGSGQTCDTGLMADVGTVCDFSR 384  
 DB 293 AAQTLRSFCWQGLNTPNSDPDHFDTALTFTQDLGSGVSTCDTGLMADVGTVCDFPAR 352  
 QY 385 CSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACASLNGVNQDS-HMVASMLSNLDHSPW 443  
 DB 353 CAIVEDDGLQSAFTAAHELGHVFNMLHDNSKPCNTNLNGQSSSRHVPVMAHVDPEPW 412  
 QY 444 SPCSGYMITFSLDNGHGECLMDKPNPIQLPGLDPTGTSYDANRQCQTFGBEDSKHCPDAA 503  
 DB 413 SPCSARFITDLDNGYGHCLLDKPEAPLHLPATPPGKDYADRCQLTFFGPDSSHCPLP 472  
 QY 504 STCSTLMTCTSGVLVCOIKHPFPAWADGTCGEGKWCINGKCNKNHKKHFDTPPHGSMG 563  
 DB 473 PPCAALWCSGHLNHAMCQTKHSPWADGTPCGSSQACMGGRCLHVDQLKDFNVPAQGWG 532  
 QY 564 MWGPWGDCSRTCGGVQVYTMRECDNVPKNGKCYCEGRVRYRSCNLEDCPDNNGKTFRE 623  
 DB 533 PWGPWGDCSRTCGGVQVFSRSDCTRPVPRNGKCYCEGRITFRSCNTENCCHGSALTFR 592  
 QY 624 EQCEAHNEFSKASFGSGPA-VEMIPKYAGVSPKDRCKLIQAKGIGYFFVLQPKVVDGTP 682  
 DB 593 EQCAAYNHRDIL-FKSPFGPMWVPRVTGVAPRQCKLTQCARALGYVYVLEPRVADGTP 651  
 QY 683 CSPDSTSVYQGVQKAGCDRIIDSKKFKDKCGVCGNGSTCKKISGVSITSAKPGYHDI 742  
 DB 652 CSPDSTSVYQGVQKAGCDRIIGSKKKFKDKMWCGGDSRCSQSGSKFKFRYGSDV 711  
 QY 743 TIPTGATNIEVKQRNQRSGRNNGSFLAIKAADGTYYILNGDYTLSTLEQDIMYKGVW-LRY 801  
 DB 712 TIPAGATHILVRQ--QGGSLGKSIYLALKLSDSGYALNGEYTLMPSTPDVVLPGLVSLRY 769  
 QY 802 SGSSAALERIRSPPLKEPTIQLVTGNALRPKIKTYTFVKK 844  
 DB 770 SGATAASETLSGHGFLAQLTLQVLVAGNPQNALRYSFVPR 812  
 RESULT 15  
 QYYS95  
 ID QYYS95 PRELIMINARY; PRT; 839 AA.  
 AC QYYS95;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE AggreCanase-1.  
 GN Name=ADAMTS-4;





**This Page Blank (uspto)**